Estudios experimentales de la enseñanza entre pares en la niñez

Efectos y moduladores en distintos dominios del conocimiento

por

Diego Pablo de la Hera

Trabajo de tesis para optar al título de Doctor en Psicología

Codirección:

Cecilia Inés Calero · Mariano Sigman

Lugar de trabajo:

Laboratorio de Neurociencia - Universidad Torcuato Di Tella

14 de diciembre de 2021

Facultad de Psicología

Universidad Nacional de Córdoba

Esta obra está bajo una licencia Creative Commons Atribución 4.0 Internacional:

https://creativecommons.org/licenses/by/4.0/deed.es

Resumen

La enseñanza es fundamental para el desarrollo y el mantenimiento de la cultura. Se cree que las personas tenemos una tendencia natural hacia esta forma de comunicación, tanto para percibirla como para impartirla. Esta inclinación natural es explotada por la tutoría entre pares, un tipo de enseñanza que promueve el aprendizaje de contenidos en diversas áreas y niveles educativos, como así también el desarrollo de competencias. Sin embargo, su potencial en algunos dominios del conocimiento relevantes para el mundo de hoy permanece sin explorar o subexplorado, y no es claro cuál es el alcance de sus efectos y bajo qué condiciones se manifiestan. En esta tesis evaluamos y caracterizamos la tutoría entre pares en la niñez para dos de estos dominios, que suelen presentar dificultades en el aprendizaje: el cambio conceptual en ciencias, y la programación de computadoras. Con este fin, presentamos tres estudios experimentales, controlados, pseudoaleatorizados, con participantes de entre 6 y 12 años de edad. En estos estudios evaluamos el impacto de la tutoría entre pares, tanto para quienes aprenden como para quienes enseñan, a corto y largo plazo; su interacción con variables relacionadas, y diferencias con otras formas de enseñanza. El procedimiento general consta de cuatro etapas secuenciales: (1) de instrucción opcional, (2) de evaluación inicial, (3) de enseñanza entre pares o revisión autónoma, y (4) de evaluación final. Los resultados muestran que la tutoría entre pares sería efectiva en los dominios explorados, identificando diferencias con otras formas de enseñanza. Además, sugieren interacciones con la metacognición, y diferencias por edad y género, entre otras variables. Por otra parte, si bien no describen beneficios de enseñar o a largo plazo, no revelan perjuicios tampoco. Asimismo, cabe destacar que nuestros estudios ofrecen también aportes metodológicos originales. En resumen, este trabajo presenta evidencia a favor del uso de la tutoría entre pares como estrategia pedagógica, subrayando su rol complementario frente a otros tipos de enseñanza, como posible promotora de formas alternativas de relación con el conocimiento.

Palabras clave

enseñanza; tutoría entre pares; niñez; cambio conceptual; programación de computadoras; creencias epistémicas; metacognición; socioconstructivismo

Abstract

Experimental studies on peer-tutoring in children: Effects and regulators in different knowledge domains.

Teaching is fundamental to the development and maintenance of culture. Humans seem to have a natural tendency towards this form of communication, both as learners and as teachers. This natural skill is leveraged by peer tutoring, a way of teaching that facilitates learning in a variety of fields and educational levels, as well as the development of skills. However, its potential in some knowledge domains relevant to our world today has not been sufficiently explored, and it is not clear which are its effects and under what conditions they become apparent. In this thesis we evaluate and characterize peer tutoring in children for two of these domains, that often present learning difficulties: conceptual change in science, and computer programming. To this end, we present three experimental, controlled, pseudo-randomized studies with participants between 6 and 12 years of age. In these studies, we evaluate the short- and long-term impact of peer tutoring on both learners and teachers, its relationship with related variables, and differences with other forms of teaching. The overall procedure consists of four sequential stages: (1) optional instruction, (2) initial assessment, (3) peer teaching or autonomous review, and (4) final evaluation. The results suggest that peer tutoring may be effective in the domains explored, identifying differences with other forms of teaching. In addition, they describe some interactions with metacognition, and age and gender differences, among other variables. On the other hand, although they do not show any teaching or long-term benefits, they do not reveal any detriments either. It is also worth highlighting that our studies offer original methodological contributions to this research field. In summary, this work provides evidence supporting peer tutoring as a pedagogical strategy, emphasizing its complementary role to other teaching approaches, as a possible promoter of alternative ways of relationship with knowledge.

Keywords

teaching; peer tutoring; childhood; conceptual change; computer programming; epistemic beliefs; metacognition; social constructivism.

Índice general

Resumen....ii
Abstract....iii
Índice general....iv
Lenguaje no sexista....x
Agradecimientos....xii
Capítulo 1: Introducción....1
1.1 Cultura y sociedad....1
1.2 Enseñanza....2
1.2.1 Definiciones de enseñanza....3
1.2.1.1 Definiciones mentalistas de enseñanza....3
1.2.1.2 Definiciones culturales de enseñanza....3
1.2.1.3 Definiciones funcionalistas de enseñanza....4
1.2.1.4 Propuesta de definición integrada....4
1.2.2 Enseñanza y acumulación cultural....5
1.2.3 Predisposición a la enseñanza....6
1.2.3.1 Sensibilidad natural....6
1.2.3.2 Ejercicio espontáneo....7
1.2.3.3 Perspectiva integrada....7
1.3 Tutoría entre pares....8
1.3.1 Historia de la tutoría entre pares....9
1.3.2 Rasgos distintivos....10
1.3.2.1 Como forma de aprendizaje asistido por pares....10
1.3.2.2 Como forma de tutoría....13
1.3.2.3 Como forma de enseñanza....13
1.3.3 Efectos y moduladores de la tutoría entre pares....14
1.3.3.1 Nivel de conocimiento....14
1.3.3.2 Teoría de la mente....15
1.3.3.3 Metacognición....16
1.3.3.4 Aspectos socioafectivos....17
1.3.4 Formas de implementación....18
1.3.5 Limitaciones....20

1.4 Estudios experimentales de la enseñanza entre pares en la niñez.....21
1.4.1 Objetivos específicos.....21
1.4.2 Hipótesis de trabajo.....22
1.4.3 Desarrollo.....22
Capítulo 2: Enseñanza entre pares y cambio conceptual.....24
2.1 Introducción.....24
2.1.1 Cambio conceptual.....24
2.1.1.1 El origen de los conceptos.....24
2.1.1.2 Procesos de cambio conceptual.....25
2.1.1.3 Cambio conceptual en astronomía observacional.....26
2.1.1.4 Importancia pedagógica del cambio conceptual.....28
2.1.2 Interacción social y construcción de conocimiento.....28
2.1.2.1 Conflicto sociocognitivo.....29
2.1.2.2 Formas de resolución del conflicto.....29
2.1.2.3 Efectos negativos de la interacción social.....31
2.2 Estudio principal 1.....31
2.2.1 Introducción.....31
2.2.2 Métodos.....33
2.2.2.1 Participantes.....33
2.2.2.2 Software y equipamiento.....33
2.2.2.3 Diseño experimental.....34
2.2.2.4 Puntaje de dimensión y nivel de conocimiento.....41
2.2.2.5 Coincidencia por temas y similitud de conocimiento.....44
2.2.2.6 Cuestionario de comportamiento.....45
2.2.2.7 Análisis de datos.....45
2.2.3 Resultados.....45
2.2.3.1 Cambios de nivel de conocimiento.....46
2.2.3.2 Cambios de similitud de conocimiento.....49
2.2.3.3 Adopción de respuestas y cambios de nivel de conocimiento.....53
2.2.3.4 Efecto en los modelos mentales de la Tierra.....55
2.2.4 Discusión.....57
2.2.4.1 La interacción aumenta los niveles de conocimiento.....57
2.2.4.2 Efecto tutor.....58
2.2.4.3 Interacción colaborativa y consenso.....60

2.2.4.4 Del pensamiento aristotélico al copernicano....61
2.2.4.5 Impacto pedagógico....62
2.3 Estudio adicional....63
2.3.1 Introducción....63
2.3.2 Métodos....64
2.3.2.1 Definiciones....64
2.3.2.2 Construcción de espacios semánticos....66
2.3.2.3 Diferencias entre subcorpus....67
2.3.3 Resultados....67
2.3.4 Discusión....68
2.4 Conclusiones....69
Capítulo 3: Enseñanza entre pares de lenguajes de programación....72
3.1 Introducción....72
3.2 Estudio principal 2....75
3.2.1 Métodos....75
3.2.1.1 Muestra....75
3.2.1.2 Plataforma de programación....76
3.2.1.3 Interfaz de usuarie....80
3.2.1.4 Videos....82
3.2.1.5 Procedimiento....83
3.2.1.6 Desempeño....87
3.2.1.7 Configuración experimental....89
3.2.1.8 Análisis de datos....90
3.2.2 Resultados....91
3.2.2.1 Efecto del tipo de instrucción....91
3.2.2.2 Efecto de enseñar....99
3.2.2.3 Efectos a largo plazo....103
3.2.3 Discusión....105
3.2.3.1 Aprendiendo de une par....105
3.2.3.2 El efecto de enseñar a une par....108
3.2.3.3 Enseñanza entre pares y metacognición....109
3.2.3.4 Diferencias entre géneros....110
3.3 Conclusiones....110
Capítulo 4: Enseñanza entre pares de reglas jerárquicas....113

4.1 Introducción....113
4.2 Estudio inicial....118
4.2.1 Métodos....118
4.2.1.1 Participantes y procedimiento....118
4.2.1.2 Clasificación de ejercicios....119
4.2.1.3 Análisis....121
4.2.2 Resultados....121
4.2.2.1 Precisión....121
4.2.2.2 Tiempo de repuesta....122
4.2.2.3 Saliencia de atributos....124
4.2.3 Discusión....126
4.3 Estudio principal 3....128
4.3.1 Métodos....129
4.3.1.1 Resumen....129
4.3.1.2 Muestra....130
4.3.1.3 División por grupos y parejas....132
4.3.1.4 Procedimiento....133
4.3.1.5 Plataforma experimental....135
4.3.1.6 Etapas....142
4.3.1.7 Metacognición....145
4.3.1.8 Análisis....147
4.3.2 Resultados....148
4.3.2.1 Efecto del tipo de instrucción....148
4.3.2.2 Efecto de enseñar....153
4.3.3 Discusión....156
4.4 Conclusiones....159
Capítulo 5: Discusión final....161
5.1 Aproximación metodológica....161
5.2 Resumen de los estudios....162
5.2.1 Cambios conceptuales....163
5.2.2 Lenguajes de programación....164
5.2.3 Reglas jerárquicas....165
5.3 Avances reportados....168
5.3.1 Dominios de relevancia....168

5.3.2 Rol pedagógico complementario....169
5.3.3 Efecto tutor....169
5.3.4 Efectos a largo plazo....170
5.3.5 Metacognición....170
5.3.5.1 Efectos sobre metacognición....170
5.3.5.2 Modulación por metacognición....171
5.3.6 Variables de contexto y organización....172
5.3.7 Aportes metodológicos....173
5.4 Direcciones futuras....173
5.4.1 Equilibrio precisión-velocidad....174
5.4.2 Efecto tutor....174
5.4.3 Metacognición....175
5.4.4 Interacciones....177
5.4.5 Validez ecológica y generalización....177
5.5 Conclusión final....179
5.5.1 Educación....179
5.5.2 Cultura y sociedad....179
Referencias....181
Glosario....207
Índice de figuras....210
Índice de tablas....211
Anexos....212
Anexo 2.1: Cuestionario completo....212
Anexo 2.2: Guía de codificación....214
Anexo 2.3: Correspondencia entre preguntas y temas....230
Anexo 2.4: Temas del Esquema de codificación 2....231
Anexo 3.1: Catálogo de imágenes y sonidos....233
Anexo 3.2: Libro de instrucciones....234
Anexo 3.3: Libro de ejemplos....238
Anexo 3.4: Eventos en videos....239
Anexo 3.5: Guión de instrucción....240
Anexo 3.6: Guión de explicación....255
Anexo 3.7: Distancias de Levenshtein....257
Anexo 3.8: Tabla de traducción distancias de Levenshtein....260

Anexo 4.1: Guión de instrucción.....262

Lenguaje no sexista

El lenguaje es “una construcción social con capacidad no sólo de refractar sino de construir realidad(es) y cultura”, como fue manifestado por el H. Consejo Superior de la Universidad Nacional de Córdoba.[1] El español posee un *masculino* no marcado (también llamado “genérico”) para referirse a grupos humanos compuestos por hombres y mujeres, o para cuando la composición de un grupo es desconocida o irrelevante; por ejemplo “los niños”. Este rasgo sería el “correlato gramatical obvio y esperable de un ordenamiento social que privilegia históricamente al hombre, un patriarcado ancestral de la especie humana”.[2]

Por esta razón, se han propuesto diversas estrategias para remediar este aspecto del lenguaje y en general manifestarse políticamente respecto del contexto sociocultural que le habría dado origen. Estas prácticas, como el desdoblamiento (“las niñas y los niños”), la paráfrasis (“la niñez”), la sustitución de vocales (“lxs niñxs”), entre otras, se conocen como lenguaje *no sexista* o *inclusivo*.[3]

En particular, el H. Consejo Superior de la Universidad Nacional de Córdoba ha recomendado a los Consejos Directivos de sus Facultades “habilitar el uso de lenguaje inclusivo en comunicaciones oficiales y administrativas, dispositivos legales, trabajos y producciones académicas, y proyectos de tesis”, con el objetivo de visibilizar “aquellas identidades que fueron/son invisibilizadas históricamente bajo la colectivización masculina”.[1] En este sentido, el H. Consejo Directivo de la Facultad de Psicología aprobó el uso de lenguaje no sexista, incluyendo “expresiones no reguladas por la academia como ‘e’ y ‘x’ en lugar de vocales que designen pertenencias de identidad de género históricamente normativizadas”.[4]

La problemática del masculino genérico, que puede hacernos suponer que una persona o conjunto de personas se encuadran dentro del género masculino[1] solamente, es de particular importancia en este trabajo, en el que estudiamos diferencias de género, y en el que además tratamos con dominios del conocimiento en los que se han observado sesgos socioculturales de base.

Por lo tanto, el presente trabajo se encuentra escrito en lenguaje inclusivo no sexista, utilizando el morfema “-e” como forma de género neutro o común [2], según

1 En este trabajo emplearemos “masculino” y “femenino” para referirnos a categorías de género.

los lineamientos propuestos en [5] y [6].[2] Brevemente, consiste en sustituir los morfemas “o” y “a”, cuando éstos indican identidad de género, en pronombres, sustantivos y adjetivos, por el morfema “e” (por ejemplo, “les niñes”). Esta práctica se diferencia de otras estrategias en que no sólo busca romper con el sexismo implícito del genérico masculino, sino también con el binario femenino/masculino, sin afectar además la economía del lenguaje y la lectura fluida.[3]

[1] H. Consejo Superior. (2019). Resolución 1095/2019. Universidad Nacional de Córdoba.

[2] Kalinowski, S., Gasparri, J., Pérez, S. I., & Moragas, F. (2020). Apuntes sobre lenguaje no sexista e inclusivo. UNR Editora, Editorial de la Universidad Nacional de Rosario.

[3] Tosi, C. L. (2019). Marcas discursivas de la diversidad. Acerca del lenguaje no sexista y la educación lingüística: Aproximaciones al caso argentino. Álabe, 20. https://doi.org/10.15645/Alabe2019.20.11

[4] H. Consejo Directivo. (2019). Resolución 347/2019. Facultad de Psicología, Universidad Nacional de Córdoba.

[5] Gómez, R. (2016). Pequeño manifiesto sobre el género neutro en castellano.

[6] Asamblea No Binarie (2018). Guía de Uso de Lenguaje Inclusivo

2 Las guías no son explícitas acerca de ciertos casos especiales. En particular, usaremos: “le(s) individue(s)”, “le(s) miembro(s)”, “le(s) modelo(s)”, “le(s) padre(s)”, le(s) ser(es) humane(s)”, “le(s) sujeto(s)”.

Agradecimientos

A mis viejes, por acompañar y apoyar todos mis proyectos, por encender y alentar la llama de la curiosidad y del idealismo en mí, por enseñarme a amar la vida. Gracias.

A mis hermanes, Martín y Rocío, por bancarme en todas.

A mis abueles, que ya no están, pero cuya obstinación, ingenio, curiosidad, y alegría, me han marcado para siempre.

A Estela, Silvia, Adri, Carmen, Negro, María Esther, Coca, Nico, Laly, Davi, y toda mi familia, de la sangre y del corazón, por la compañía, el apoyo y los consejos.

A mis directores, Ceci y Mariano. Por confiar en mí, por abrirme las puertas del laboratorio, por apoyarme, guiarme, alentarme y aguantarme, y por permitirme transitar una de las experiencias más enriquecedoras de mi vida.

A mis colegas y amigues del Laboratorio de Neurociencia. Por inspirarme constantemente y ayudarme a mejorar mi trabajo, con pasión, espíritu crítico y buena onda. Por los almuerzos, los seminarios, las charlas, los debates, y una curiosidad infinita. Gracias Agus, Alejo, Analí, Andre, Andrés, Andy, Ari, Caro, Ceci, Chu, Diego, los Fede, Guada, Ioni, Joaco, las Juli, Lean, Lu, Mariano, Martina, Mati, Meli, Tami, y Vero.

A mis querides amigues, Brunito, Gusti, Santi, Javi, Jime, Romi, Maru, y Sepu.

A Tomi, compañero de alegrías y tristezas, de aventuras y charlas. Por su amistad, afecto y generosidad incondicionales.

A Anna, por su magia contagiosa, por bancarme, por las risas, por las lágrimas.

A Danel, Luli, Javi, Maru, Fer, Belén, Martín, y todes mis amigues, familia, y amigues que son familia, en el Valle de Calamuchita, que me acompañaron, apoyaron y bancaron en los últimos meses.

A Mati, por acompañarme a conocer nuevos mundos. Por estar y escucharme en los primeros años de esta aventura.

A Gabi, por inspirarme y animarme a recorrer caminos inciertos.

A Pato, por la música, y la compañía generosa, paciente y comprensiva en la larga recta final.

A mis colegas y amigues bioquímiques, Euge, Juli, Pehuén, Meli, y Andre.

A Martín, Juan, y mis compañeres y amigues del Subsuelo Cultural y del Ensamble Vocal Di Tella, personas mágicas que me enseñaron mundos maravillosos y sin quienes estos años no habrían sido los mismos.

A Scann, por inspirarme y empujarme a explorar nuevos horizontes.

A Santi, por invitarme a la charla donde empezó todo.

A Cande, compañera en los inicios, madrugados y remotos, de esta aventura.

A Maxi, por su compañía y escucha en tantos almuerzos de la Di Tella.

A Belén y Luciana, que con afecto, paciencia y compromiso, y entre charlas y risas, me acompañaron en las escuelas y me ayudaron a mejorar los procedimientos.

A Tami, Anto, y todes les técniques y asistentes, que bancaron mi inexperiencia como jefe y que con profesionalismo y dedicación sacaron los estudios adelante.

A les docentes, directives, personal no docente, padres y niñes de las escuelas, que confiaron en nosotres, nos abrieron las puertas, y dieron todo para que nuestro trabajo fuera un éxito.

A les niñes y adultes que participaron de nuestros estudios y pilotos.

A Nicolás Roulet, Anahí Rosso, Juan Kamienkowski y Alejo Salles, por sus contribuciones a los estudios complementarios de los capítulos 2 y 4.

A Ariel Haimovici y Julieta Goldstein, por sus aportes al estudio principal del Capítulo 4.

A Soledad Segretin, por sus valiosas contribuciones como evaluadora externa.

A Ceci y Andre, por invitarme a participar de la organización de la LASchool 2016, una experiencia extenuante y enriquecedora, que me abrió puertas al mundo. Y al equipo de la LASchool 2018, por aceptar mi participación, esta vez como estudiante.

A mi amiga y colega Azzurra, que confió en mí para dirigir un proyecto que me hizo crecer muchísimo y que me ayudó a convencerme de que estaba listo para doctorarme.

A Cansu, Daniil, y colegas del *Max-Planck-Institut für Bildungsforschung*, compañeres de esa montaña rusa pre-pandémica.

A Pitu, compañeres del Ensamble Vocal Encanto, y amigues en Berlín, que me acompañaron y sostuvieron en esos locos tres meses.

A Hugo, Feli, Ger, Lu, Mariela, Clara, Nico, y colegas del IQUIFIB, que me ayudaron a dar mis primeros pasos en este mundo.

A Alexandra Elbakyan, sin cuyo osado compromiso por la ciencia abierta este trabajo habría sido mucho más difícil e incompleto.

A les revisores anónimes de nuestros trabajos publicados, cuyos comentarios nos ayudaron a mejorarlos.

Al CONICET, por la beca que hizo posible este proyecto.

A la Facultad de Psicología de la Universidad Nacional de Córdoba, por abrirme las puertas a pesar de la distancia geográfica y académica.

A la Universidad Torcuato Di Tella, por el lugar de trabajo, los recursos, y por integrarme a su comunidad.

A IBRO, por ayudarme a comunicar mi trabajo del otro lado del charco.

A la DAAD, por ayudarme a concretar el sueño de vivir y trabajar en otra ciudad.

A Educación Tecnológica, por facilitar recursos y asesoramiento que nos ayudaron a realizar el estudio del Capítulo 3.

A la Fundación y comunidad Wikimedia, por confiar en mis propuestas y ayudarme a que escribir esta tesis en pandemia fuese económicamente sostenible.

A les docentes, de todos los niveles, que con pasión apoyan y alientan la curiosidad de sus estudiantes.

A la bici, que estuvo ahí en todas todas.

Al Bicho, por la compañía incondicional.

Al universo y a la vida, por las personas que durante años, meses, días, o tan solo minutos, han tejido sus hilos con los míos, en el vasto tapiz de la cultura humana.

Capítulo 1: Introducción

En poco tiempo la especie humana ha evolucionado de un conjunto de comunidades relativamente aisladas y pequeñas a una comunidad crecientemente global. En este contexto, las relaciones interpersonales se han ido complejizando, como así también los fenómenos sociales que emergen de ellas, demandando su estudio y comprensión. La cultura es uno de estos fenómenos emergentes, y la enseñanza una de las formas en que se dan estas relaciones.

En esta introducción comenzaremos hablando sobre la cultura y sus dinámicas de construcción y acumulación. Luego nos centraremos en la enseñanza, como instrumento para su desarrollo y mantenimiento. Discutiremos definiciones, y teorías acerca de la manifestación temprana de este proceso.

A continuación, nos enfocaremos en un tipo de enseñanza en particular: la tutoría entre pares. Discutiremos su historia, rasgos característicos, beneficios, variables de las que depende o *moduladores*, formas de implementación y limitaciones. Finalmente, presentaremos los objetivos de este trabajo, en relación a los problemas y limitaciones planteados, resumiendo las investigaciones que llevamos adelante para abordarlos, y que desarrollaremos en los próximos capítulos.

1.1 Cultura y sociedad

La humanidad es una especie cultural; su entorno es en parte un fenómeno social, constituido por interacciones entre individues y artefactos culturales (Vygotsky, 1978). Entendemos como *cultura* al conjunto complejo que incluye al conocimiento, las creencias, el arte, la moral, el derecho, las costumbres y cualquier otra capacidad y hábito que las personas adquieren como miembros de una sociedad (Tylor, 1974).

Sin embargo, la cultura no es estática, sino que se construye y afecta colectivamente de forma constante. Según la teoría del *construccionismo social*, construimos entendimientos sobre el mundo en conjunto, desarrollando significados en coordinación con otres, que forman la base de las suposiciones compartidas sobre la realidad (Littlejohn & Foss, 2009). Así, podemos imaginar a la sociedad como una red cuyos nodos están ocupados arbitrariamente por individues. Cada une de elles obtiene información del contexto sociocultural, la interpreta según el conjunto único

de sus experiencias pasadas (Mead et al., 2015) –una especie de huella digital interpretativa–, y eventualmente la devuelve, modificada, hacia el mismo contexto.

Pero para poder comprender este fenómeno en toda su magnitud, es necesario considerar también su dimensión temporal. Es decir, la cultura no es sólo el producto de la construcción constante y colaborativa, sino también de la acumulación en el tiempo de los logros culturales de quienes nos precedieron (Boyd et al., 2011), de forma análoga a cómo la información genética es el resultado de milenios de evolución. De este modo las interacciones sociales forman parte fundamental de la ontogenia humana en este contexto sociocultural, preservando los logros culturales generación tras generación y superando cualquier otro proceso observado en otras especies (Stöber, 2016).

Si bien parte de esta información cultural puede adquirirse por experiencia directa, muchos de estos conocimientos son cognitivamente opacos. Es decir, no son intuitivos ni obvios, sino el resultado de complejas construcciones sociales. ¿Cómo es posible, entonces, que podamos transmitir este conocimiento de forma eficiente y fiel para ser perpetuado a lo largo de la historia? ¿Y quiénes y cómo participamos en estos procesos de acumulación cultural?

1.2 Enseñanza

Aprendemos como resultado de las experiencias que nos toca vivir. En algunos casos aprendemos a solas. Por ejemplo, cuando inferimos regularidades del mundo físico que nos rodea, observándolo. En otros casos, aprendemos socialmente, con otres (*aprendizaje social*). Sin embargo, a veces el aprendizaje puede presentar dificultades o ser ineficiente, y es conveniente o incluso necesario la participación de un facilitador. Este facilitador es la *enseñanza*.

Las ciencias cognitivas se han abocado con gran énfasis al estudio del aprendizaje. Por el contrario, existe un llamativo vacío en la investigación de la capacidad de enseñar (Strauss et al., 2014), largamente ignorada también en el estudio del comportamiento animal (Thornton & Raihani, 2008). ¿Pero qué entendemos exactamente por “enseñar”?

1.2.1 Definiciones de enseñanza

La enseñanza ha sido definida de distintas maneras según el campo desde el que se la considere. Esto plantea dificultades para un abordaje interdisciplinario. Por este motivo, ponernos de acuerdo en a qué nos referimos con “enseñar” es importante para facilitar el diálogo entre campos y permitir que se nutran unos a otros.

Una posible clasificación de las distintas formas de definir enseñanza ha sido propuesta por Kline (2014), comprendiendo tres familias de definiciones: mentalista, cultural, y funcionalista.

1.2.1.1 Definiciones mentalistas de enseñanza

Según las definiciones *mentalistas* de la enseñanza, entendemos por “enseñanza” a todo comportamiento que tiene la *intención* de facilitar el aprendizaje en une otre, es decir de reducir la diferencia de conocimiento con aquelle. Para que exista esa *intención* por parte de quien enseña, es necesario que éste sea capaz de identificar dicha diferencia. Es decir, requeriría ciertas habilidades psicológicas, como la de construir representaciones mentales de une otre o Teoría de la Mente. Así, estas definiciones se enfocan en diferencias en estas habilidades psicológicas, entre especies o individues, para explicar variaciones en su supuesta habilidad para enseñar, excluyendo incluso aquelles que, por su naturaleza o grado de desarrollo, carecerían de las habilidades necesarias.

Un ejemplo de este tipo de definición es la que ofrecen Strauss y Ziv (2012), que la presentan como “una forma particular de *comunicación social*, guiada por la intención de *causar aprendizaje* en les otres”, como desarrollaremos más adelante.

1.2.1.2 Definiciones culturales de enseñanza

Las definiciones culturales de enseñanza comprenden el conjunto muy limitado de comportamientos, explícitamente reconocidos como “enseñanza” por todes les participantes, que se dan en el contexto específico de las aulas en las sociedades occidentales y que buscan forzar el aprendizaje en estudiantes pasives y posiblemente desinteresades. Estas definiciones se enfocan en la comparación con otras formas de aprendizaje social e individual fuera del contexto cultural occidental.

1.2.1.3 Definiciones funcionalistas de enseñanza

A diferencia de las definiciones mentalistas, basadas en la psicología de quien enseña, y culturales, basadas en el contexto de enseñanza, las definiciones funcionalistas se enfocan en las raíces evolutivas y el diseño adaptativo de enseñar. Es decir, entienden como *enseñanza* al conjunto de comportamientos seleccionados evolutivamente por facilitar el aprendizaje en otres. Algunas de estas definiciones postulan que estos comportamientos podrían haber sido seleccionados por aumentar la precisión del aprendizaje en procesos de aprendizaje social preexistentes, dando origen a distintos perfiles de comportamiento o tipos de enseñanza (Hoppitt et al., 2008). Es decir, según estas definiciones, los comportamientos que entendemos por enseñanza se habrían seleccionado por facilitar procesos de aprendizaje social que de otro modo habrían sido menos eficientes, menos precisos, o imposibles.

Un ejemplo de este tipo de definición es la definición operacional de enseñanza de Caro y Hauser (1992), para distinguirla de otras formas de aprendizaje social, según la cual hablamos de enseñanza si el comportamiento: (1) depende de la presencia de une aprendiz inexperte; (2) no ofrece beneficios inmediatos (y en algunos casos puede representar costos) para quien enseña; y (3) se puede demostrar que causa que le aprendiz aprenda algo antes o más eficientemente de lo que habría ocurrido de otro modo, o que simplemente no habría ocurrido.

Otro ejemplo es el que ofrecen Csibra y Gergely (2009) en su teoría de la *Pedagogía Natural*, entendida como un conjunto de adaptaciones (en quien enseña y en quien aprende) para facilitar el aprendizaje de une otre, para un tipo de contenido en particular (conocimiento generalizable) y para un tipo de aprendizaje en particular (transferencia de conocimiento), en una especie en particular (humana), como se verá más adelante.

1.2.1.4 Propuesta de definición integrada

Kline (2014) propone una taxonomía de enseñanza, basada en la aproximación funcionalista, pero que busca ofrecer un marco común a lo ancho de estas tres aproximaciones mayoritarias, en el que cada una podría nutrirse de la mirada de las otras dos. Por ejemplo, la integración de definiciones funcionalistas (centradas en problemas adaptativos y adaptaciones comportamentales) y definiciones mentalistas

(enfocadas en mecanismos psicológicos que determinan el comportamiento) podría impulsar la investigación de circunstancias socioecológicas bajo las cuales cierto rasgo psicológico podría haber sido seleccionado (Kline, 2014).

Para aprender hace falta que le aprendiz (a) ponga atención, y (b) tenga acceso, a información relevante, entendida como información novedosa, de utilidad para le aprendiz, y de algún modo conectada a su conocimiento previo (Sperber & Wilson, 1996). Según la definición integrada propuesta, entendemos como enseñanza al conjunto de adaptaciones del comportamiento, tal vez derivadas de comportamientos de aprendizaje social preexistentes, que hacen al aprendizaje posible, más eficiente, o más preciso, cuando la atención o el acceso a la información, o ambos, escasean.

Basado en problemas adaptativos específicos inherentes al aprendizaje social, caracterizados por distintas combinaciones de atención y acceso a la información, Kline postula una taxonomía de cinco tipos de enseñanza. Por un lado, los tipos de enseñanza (1) *por tolerancia social* y (2) *por provisión de oportunidades* ofrecerían ventajas adaptativas en casos donde la atención es suficiente pero el acceso a la información está comprometido. Concretamente, cuando (1) la información depende de la observación de une modelo, y (2) cuando depende de experiencias raras o peligrosas, respectivamente. Por otra parte, los tipos de enseñanza (3) *por refuerzo del estímulo* y (4) *por retroalimentación evaluativa* habrían evolucionado en contextos de acceso a la información adecuado, y escasa atención. Puntualmente, cuando falla la atención (3) al estímulo, o (4) a las consecuencias del comportamiento. Por último, el tipo de enseñanza (5) *activa directa* habría sido seleccionado en contextos donde tanto la atención como el acceso a la información son insuficientes.

1.2.2 Enseñanza y acumulación cultural

Como mencionamos al comienzo, la especie humana habría evolucionado en un nicho cultural, de modo tal que depende en gran medida de adaptaciones culturales acumulativas, demasiado complejas para ser desarrolladas de forma individual, que habrían coevolucionado con capacidades culturales (Boyd et al., 2011; Kline, 2014). La enseñanza *activa directa*, mencionada en la sección anterior, podría ser una de estas capacidades.

El conocimiento cultural es considerablemente arbitrario y cognitivamente opaco (Csibra & Gergely, 2009). Por esta razón, la enseñanza podría haber coevolucionado con la cultura por su capacidad catalizadora de procesos de aprendizaje que de otra manera serían menos eficientes, propensos a errores, o simplemente imposibles (Kline, 2014). Esta afirmación ha sido apoyada experimentalmente (Dean et al., 2012; Kurzban & Barrett, 2012), aunque habría evidencia también de que no siempre es necesaria y otras vías de acumulación cultural existirían también (Caldwell & Millen, 2009).

¿Pero nos enseñamos a enseñar? ¿O estamos naturalmente predispuestos a la enseñanza?

1.2.3 Predisposición a la enseñanza

A lo largo de los últimos años, ha tomado relevancia la idea de que les seres humanes estaríamos preparades para participar de los procesos de enseñanza-aprendizaje desde edades tempranas. Esta supuesta predisposición a la enseñanza puede entenderse de dos maneras: como sensibilidad natural (a ser enseñade), y como ejercicio espontáneo (de enseñar). A continuación, desarrollaremos ambas perspectivas y presentaremos también una posible integración.

1.2.3.1 Sensibilidad natural

Por un lado, la teoría de la Pedagogía Natural de Csibra y Gergely (2009) se enfoca en la *receptividad* temprana a la enseñanza. Esta definición funcionalista entiende a la enseñanza como una forma de comunicación que habría sido seleccionada evolutivamente por su capacidad de facilitar la transmisión de conocimiento abstracto, genérico y cognitivamente opaco, que de otra manera sería difícil de adquirir. Este protocolo de comunicación se caracteriza por demostraciones ostensivas, es decir señales no verbales que le emisore emplea para señalar selectivamente la información que debe ser adquirida y generalizada por le receptore. De este modo, acelera el ritmo al que aprendices humanes inexpertes obtienen información confiable, novedosa y relevante acerca del mundo.

Según estos autores, les infantes están preparades para situarse del lado receptor de la Pedagogía Natural visto que: 1) son sensibles a claves ostensivas; 2) esperan un referente u objeto aludido en contextos de comunicación ostensivos; y 3) están

sesgades a interpretar el contenido de la comunicación ostensiva como generalizable, es decir no limitado a hechos episódicos.

Al ser sensibles a la enseñanza, les niñes estarían naturalmente preparades para adquirir los artefactos culturales transmitidos de esta manera en las diversas actividades estructuradas por adultes de las que participan (Calero et al., 2018). Además, considerando la naturaleza bidireccional de los procesos de enseñanza (Vygotsky, 1978), ajustados constantemente según la comprensión manifestada por le aprendiz, y la interpretación que le maestre hace de ésta (Strauss et al., 2014), cabe suponer que les niñes afectarían el desarrollo de los saberes culturales en el momento mismo en el que les son transmitidos.

1.2.3.2 Ejercicio espontáneo

Por otra parte, la capacidad de enseñar pareciera surgir de forma universal e invariable siguiendo una ruta de desarrollo (Calero et al., 2018), lo cual ha llevado a postular que podría tratarse de una habilidad natural (Calero et al., 2015; Strauss et al., 2002, 2014; Tomasello et al., 2007). En particular, Strauss y Ziv (2012) postularon que esto sería así ya que, a pesar de su complejidad, la enseñanza aparece desde la infancia temprana sin instrucción aparente

Si bien la definición mentalista de estes autores subraya la necesidad de teoría de la mente para enseñar, otras miradas similares aceptan la participación de capacidades de lectura de la mente implícitas y automáticas, admitiendo la participación de niñes aún menores (Calero et al., 2018). Esta capacidad temprana de enseñar daría a les niñes un rol activo en la transmisión cultural, al menos después de los primeros dos años de vida, antes de lo cual el grado de adoctrinamiento cultural podría ser demasiado limitado (Stöber, 2016).

1.2.3.3 Perspectiva integrada

Hasta aquí vimos, por un lado, que les niñes son sensibles al protocolo de comunicación especializado para la enseñanza, postulado por la teoría de la Pedagogía Natural; y, por otro lado, que son capaces de enseñar desde edades tempranas, según la teoría de la enseñanza como habilidad cognitiva natural. ¿Pueden, acaso, enseñar empleando elles también este protocolo optimizado para la enseñanza?

En nuestro laboratorio se decidió investigar la ontogenia de la capacidad de enseñar en humanes, es decir, de la capacidad de les niñes de situarse del lado emisor de la pedagogía natural. Se observó que niñes a partir de los 3 años despliegan un amplio repertorio de claves ostensivas durante momentos de relevancia pedagógica, concluyendo que son capaces de identificar las partes relevantes de la información, y el uso adecuado de los canales necesarios para transmitirlas (Calero et al., 2015, 2018). Es decir, manifestarían una comprensión intuitiva, probablemente una habilidad natural moldeada por la interacción social, de cómo transmitir información eficientemente.

En resumen, les niñes no sólo están naturalmente preparades para aprender, sino que además están naturalmente preparades para enseñar. Esto ha sido encauzado institucionalmente en estrategias pedagógicas que explotan esta habilidad, en las que nos enfocaremos a continuación.

1.3 Tutoría entre pares

Durante la niñez, gran parte de las experiencias e interacciones están dictadas por les adultes a cargo. Sin embargo, no debemos desestimar las interacciones, intercambios de información y construcción de conocimientos que se dan entre niñes. Esta dualidad par-adulto –los “dos mundos sociales del niño”, según Piaget *(1932)*– sería una de las más comunes y claras dicotomías en los estudios del desarrollo (Bronfenbrenner, 1979). En una escuela, por ejemplo, ¿como es posible suponer que la mayoría de los intercambios dependan de les adultes, en una población conformada mayoritariamente por niñes?

La tutoría entre pares ha sido extensamente estudiada, demostrando ser útil en un amplio rango de dominios del conocimiento y de edades. Es una práctica que ofrece contextos de aprendizaje particulares para todas las partes involucradas, observándose beneficios tanto para les tutores como para les aprendices, y reportando ganancias sociales y emocionales sin costo adicional (Kalkowski, 1995; K. J. Topping, 2020). Es una de las estrategias de enseñanza con mejor relación costo-beneficio (Levin et al., 1987), que además explota la marcada asimetría en el número de docentes y estudiantes característica en la escuela de sociedades

occidentales, y que ocurre independientemente de que sea o no promovida o facilitada explícitamente.

En esta sección comenzamos revisando brevemente la historia de la aplicación de esta estrategia pedagógica. A continuación mencionaremos los rasgos característicos de esta práctica, considerando a la tutoría entre pares desde tres puntos de vista principales: como forma de aprendizaje entre pares, como forma de tutoría, y como práctica de enseñanza, cada uno con un sesgo propio hacia le aprendiz, le tutore, o ambes. Luego, revisaremos lo que sabemos acerca de los efectos y moduladores de la tutoría entre pares, teniendo en cuenta los rasgos característicos enumerados anteriormente. Habiendo mencionado los moduladores, estaremos en mejores condiciones de hablar sobre distintas estrategias de aplicación y las dimensiones a lo largo de las que son clasificadas. Y finalmente discutiremos problemas y áreas de vacancia, en cuyo marco se inscriben los objetivos de esta tesis.

1.3.1 Historia de la tutoría entre pares

Distintes autores sitúan los orígenes de la tutoría entre pares en diferentes momentos históricos (Kalkowski, 1995). Algunes entienden a la instrucción por tutores, incluyendo la tutoría entre pares, como la primera forma de pedagogía en las sociedades primitivas (Jenkins & Jenkins, 1987). Otres rastrean sus orígenes (en la civilización occidental) hasta la antigua Grecia (Wagner, 1990), citando como ejemplo a Platón y sus estudiantes (K. J. Topping, 2020).

No obstante, el antecedente histórico más evidente son los sistemas de enseñanza mediada por estudiantes avanzades o monitores. Si bien habría registros del uso de monitores en las *grammar schools* de la época isabelina del siglo XVI (K. J. Topping, 1988), éstes habrían tenido un rol accesorio (López & Narodowski, 1999). El primer uso sistemático de monitores o pares tutores se atribuye a los trabajos de Andrew Bell en el siglo XVIII, en Madras, India (Bell, 1807; López & Narodowski, 1999; K. J. Topping, 1988). Este método de enseñanza mutua o *monitoreal* se basa en la transmisión lineal de conocimiento, siguiendo una jerarquía piramidal. En él, une docente enseña a un grupo de estudiantes avanzades o monitores, quienes a su vez están a cargo de enseñar a un grupo de estudiantes cada une, alcanzando en conjunto a hasta 500 estudiantes en total (López & Narodowski, 1999). El método

fue desarrollado y difundido por Joseph Lancaster en el Reino Unido, para suplir la carencia de profesores de principios del siglo XIX (K. J. Topping, 2020), por lo que se lo conoce también como método *lancasteriano*.

En Argentina, este “último grito de la moda pedagógica europea” fue masivamente introducido en la provincia de Buenos Aires por James Thomson. De esta forma se logró establecer un proceso escolar uniforme, contrario a la heterogeneidad de los primeros años posteriores a la emancipación política de nuestro país, que además permitió al Estado proveer de enseñanza elemental a la totalidad de la población infantil con un mínimo de recursos (López & Narodowski, 1999).

En la actualidad, la tutoría entre pares ya no se usa, como en el método de enseñanza *monitoreal*, para suplir una carencia de profesores capacitades, sino más bien como herramienta para mejorar el aprendizaje, incorporada como práctica complementaria a la enseñanza tradicional, en particular en relación con dominios del conocimiento que presentan dificultades en la enseñanza y el aprendizaje (K. J. Topping, 2020). De esta manera, el interés se ha centrado en la participación de tutores cuyas capacidades sean más cercanas a las de la persona que requiere apoyo, de modo que ambes encuentren suficiente estímulo intelectual en las actividades que comparten y se beneficien de la interacción (K. Topping, 2015).

1.3.2 Rasgos distintivos

Existen estudios cualitativos acerca de las características singulares de las tutorías entre pares (ej., Dangwal & Kapur, 2009; Howe & Recchia, 2005; LeBlanc & Bearison, 2004); y otros (minoritarios) que las examinan desde una perspectiva cuantitativa-experimental (Graesser et al., 1995; Lombrozo, 2006; Roscoe & Chi, 2008). Podemos considerar los rasgos distintivos de la tutoría entre pares desde tres perspectivas: considerándola como forma de aprendizaje asistido por pares, como forma de tutoría, y como forma de enseñanza.

1.3.2.1 Como forma de aprendizaje asistido por pares

La tutoría entre pares es una forma de aprendizaje asistido por pares. Este tipo de aprendizajes involucra personas de grupos sociales similares (que habitan en una misma región geográfica, de edad y nivel socioeconómico semejantes), con niveles de conocimiento similares pero diferentes, que no son docentes profesionales, que

se ayudan activamente unes a otres a aprender, y que aprenden enseñando (Damon & Phelps, 1989; K. Topping, 2015; K. Topping & Ehly, 1998; K. J. Topping, 2020). Una forma de aprendizaje asistido por pares es el aprendizaje cooperativo, que tiene lugar en grupos pequeños, a menudo con miembros de habilidades diversas, que trabajan para llegar a un consenso en un problema en particular (Kyndt et al., 2013). Otra forma es la tutoría entre pares, que habitualmente involucra dos participantes, une más avanzade en el rol de tutore, y otre menos avanzade en el rol de aprendiz (Roscoe & Chi, 2007).

A rasgos generales, se caracterizan por y nutren del conflicto y la construcción colaborativa (K. J. Topping, 2020), beneficiándose también de condiciones particulares en materia de compromiso, individualización y personalización, comunicación, habilidades sociales, y emociones (K. J. Topping, 2020). Además, las interacciones entre niñes ofrecen oportunidades para el aprendizaje social que son cualitativamente diferentes a las que se dan en interacciones estructuradas con adultes (Stöber, 2016). De este modo, el aprendizaje asistido por pares se libera de algunos problemas que presenta la enseñanza tradicional. Por ejemplo, aunque aprender de une maestre experte puede promover aprendizajes rápidos y eficientes, también puede desalentar la exploración y el descubrimiento espontáneos (Bonawitz et al., 2011).

Concretamente, podemos considerar los rasgos típicos del aprendizaje asistido por pares a través de tres lentes: la de la interacción social, la de la credibilidad de le modelo, y la de la seguridad del entorno de aprendizaje.

Interacción social

El aprendizaje asistido por pares es, en primer lugar, una forma de interacción social. Como tal, ofrece experiencias de sociabilización, que involucran habilidades sociales y comunicacionales, con posibles efectos en la motivación y el desempeño académico de les estudiantes (Rohrbeck et al., 2003). Además, como actividad de atención conjunta, con objetivos e intenciones compartidas, y acciones coordinadas con otres, podría estimular también el desarrollo de habilidades como la Teoría de la Mente (Calero et al., 2018).

Credibilidad de le modelo

Por otra parte, el aprendizaje asistido por pares se caracteriza también por la credibilidad de le modelo de aprendizaje. Como les participantes tienen edades y contextos similares, tienen sus formas de comunicarse, y pueden referir a ejemplos y experiencias relevantes en su vida cotidiana (Kalkowski, 1995). De esta manera, les aprendices podrían comprender las explicaciones de sus tutores más fácilmente, permitiendo un aprendizaje más personalizado y efectivo (K. J. Topping, 2020), dentro de su zona de desarrollo proximal (Calero et al., 2018; Vygotsky, 1978).

Esta proximidad entre tutore y aprendiz favorece el aprendizaje y la comprensión del contenido que pretende ser enseñado. Además, esta proximidad con le tutore podría promover también el desarrollo de procesos de pensamiento y habilidades metacognitivas, que les tutores podrían modelar con el objetivo de contribuir a la comprensión por parte de sus estudiantes (Calero et al., 2018).

Seguridad del entorno de aprendizaje

Finalmente, el aprendizaje asistido por pares se caracteriza por el entorno de aprendizaje que ofrece. La confianza que les estudiantes tienen en le maestre es una parte importante pero poco estudiada del éxito de une maestre (Corriveau & Winters, 2019). En este sentido, se ha sugerido que la tutoría entre pares es cualitativamente diferente de la interacción asimétrica entre la figura de autoridad de une maestre profesional y une niñe (Piaget, 1932, 1950), lo cual implicaría distintas ventajas y desventajas.

Es decir, al involucrar niveles de conocimiento y contextos similares, las tutorías entre pares son vistas como *entornos de aprendizaje seguros*, que reducen la angustia (K. J. Topping, 2020), y en los que se establecen conversaciones de elevada mutualidad en las que le aprendiz se siente más libre para expresar opiniones, hacer preguntas, o arriesgar soluciones (Damon & Phelps, 1989). Por ejemplo, en tareas de razonamiento científico, les amigues son más propensos que les conocides a evaluar sus soluciones, justificar sus propuestas espontáneamente, elaborar y criticar el razonamiento de le otre, e involucrarse en discusiones transaccionales (Azmitia & Montgomery, 1993). Además, el sentimiento de lealtad y responsabilidad mutuas mantiene a ambas partes motivadas y centradas en la tarea (K. Topping, 2015).

1.3.2.2 Como forma de tutoría

La segunda perspectiva desde la que podemos considerar a la tutoría entre pares es como una forma particular de tutoría. En el contexto de la educación formal, las tutorías son un tipo de instrucción que involucra une docente por alumne (sean éstes pares o no) y que han demostrado ser más efectivas que otras formas de enseñanza tradicional con une docente frente a un número mucho mayor de alumnes (Bloom, 1984; P. A. Cohen et al., 1982). En las tutorías, le tutore hace preguntas que promueven procesos de pensamiento en le aprendiz, ofrece explicaciones, y corrige errores (Graesser et al., 1995; Roscoe & Chi, 2008), de forma individual y personalizada. Por tanto, cabe esperar que desde la perspectiva de le aprendiz la tutoría entre pares (una forma especial de tutoría en que aprendiz y tutore son pares) herede algunas de las características propias de estas formas de enseñanza une a une.

1.3.2.3 Como forma de enseñanza

Finalmente, la tutoría entre pares es también una forma de enseñanza. La enseñanza no sólo facilita el aprendizaje de quien es enseñade, sino que también reforzaría el aprendizaje de quien enseña (Duran, 2016; Hermida et al., 2021). Beneficios similares han sido reportados en la explicación a une misme o a une otre (Chi et al., 1994; Duran, 2016), como los beneficios de la resolución de problemas en voz alta (Lochhead & Whimbey, 1987), o la identificación de errores en software usando agentes inanimados (Hunt & Thomas, 1999; Phillips et al., 2013). Además, se ha postulado que la enseñanza podría motorizar el desarrollo de habilidades cognitivas, como funciones ejecutivas, teoría de la mente y metacognición (Calero et al., 2018). Y se la ha propuesto como herramienta para evaluar el conocimiento de le tutore, alternativa a la justificación frente a une docente experte (Calero, 2020).

Sin embargo, no todas las formas de enseñanza serían efectivas para promover aprendizaje. Por ejemplo, les docentes recurren a veces a formas improductivas de *narrar* conocimiento (*knowledge-telling*) en las que resumen la información superficialmente, a diferencia de formas productivas de *construir conocimiento* (*knowledge-building*) en las que integran y generan conocimiento mediante el razonamiento (Roscoe & Chi, 2007).

El mecanismo que subyace los beneficios de enseñar permanece en gran medida desconocido. No obstante, podemos considerar tres perspectivas. En primer lugar, la enseñanza es un proceso de interacción social bidireccional, en la que la información no fluye únicamente de le tutore a le aprendiz, sino en sentido inverso también (LeBlanc & Bearison, 2004; Vygotsky, 1978). En segundo lugar, podrían estar implicados procesos involucrados durante el acto de enseñar propiamente dicho, como mecanismos metacognitivos (Roscoe & Chi, 2008), o la recuperación de memoria de los contenidos a enseñar (Koh et al., 2018). Finalmente, hay quienes postulan que estos beneficios estarían mediados también por el hecho de prepararse para enseñar, más que por el acto de enseñar en sí (Fiorella & Mayer, 2013; Gregory et al., 2011; Marno et al., 2021; Nestojko et al., 2014), implicando por ejemplo estrategias de estudio diferentes (Bargh & Schul, 1980).

1.3.3 Efectos y moduladores de la tutoría entre pares

Muchos años de investigación en tutoría entre pares han identificado una serie de beneficios, tanto sobre el aprendizaje del tema que se pretende enseñar como sobre el desarrollo de competencias cognitivas y socioafectivas. Como veremos a continuación, se han reportado beneficios para aprendices y tutores, algunos de ellos tanto de forma inmediata como a largo plazo. Comentaremos también la relación inversa, es decir lo que se sabe acerca del rol modulador de estas variables sobre la tutoría entre pares.

1.3.3.1 Nivel de conocimiento

Nos referiremos aquí al grado de experiencia, comprensión y capacidad de aprendices y tutores acerca del tema de estudio, incluyendo también la brecha o diferencia entre ambes.

Efectos sobre nivel de conocimiento

Existe consenso de que la tutoría entre pares efectivamente facilita el aprendizaje de quienes son enseñades por une par, no sólo en términos del desempeño, sino también del entendimiento, la profundidad del aprendizaje, y el pensamiento abstracto (K. J. Topping, 2020). De esta manera ayuda a reducir la brecha entre aprendices y tutores, en una amplia variedad de dominios del conocimiento (de la Hera et al., 2019; K. Topping, 2017).

Por otra parte, esta práctica ha demostrado beneficiar no sólo a les aprendices, sino también a les tutores (P. A. Cohen et al., 1982; Duran, 2016; Goodlad & Hirst, 1989; Roscoe & Chi, 2007, 2008), y en algunos casos el beneficio habría sido incluso mayor que el observado para les aprendices (K. J. Topping, 2020).

Por último, cabe destacar también que estos efectos sobre el aprendizaje se han verificado no sólo inmediatamente, sino también a largo plazo (Ellis et al., 1998; Semb et al., 1993).

Modulación por nivel de conocimiento

Es probable que la presunción del antiguo método monitoreal presentado arriba, de que les estudiantes más avanzades debían ser les tutores, sea equivocada. Por un lado, une tutore demasiado avanzade podría dejar de ser une modelo creíble para le aprendiz, aspecto clave de la tutoría entre pares como comentamos en el apartado *Credibilidad de le modelo* de la sección “1.3.2 Rasgos distintivos”. Como consecuencia de esto, podrían verse atenuados o anulados los beneficios que la tutoría entre pares ofrece para les aprendices (K. J. Topping & Ehly, 2001).

Por otra parte, una brecha muy amplia entre tutore y aprendiz podría reducir los estímulos cognitivos para les tutores, afectando así los beneficios que la tutoría entre pares ofrece para éstes también (K. Topping, 2015; K. J. Topping & Ehly, 2001).

1.3.3.2 Teoría de la mente

Entendemos por teoría de la mente a la capacidad de atribuir estados mentales a une otre, incluyendo creencias, deseos y conocimientos, y de comprender que éstos podrían ser distintos de los propios (Premack & Woodruff, 1978).

Efectos sobre Teoría de la mente

La tutoría entre pares, como actividad de atención conjunta en la que aprendices comparten objetivos e intenciones y coordinan acciones con les tutores, podría estimular el desarrollo de la teoría de la mente en ambes participantes (Calero et al., 2018), como se comentó en el apartado *Interacción social* de la sección “1.3.2 Rasgos distintivos”.

Además, se ha postulado que la enseñanza podría ser un motor para el desarrollo de la teoría de la mente, entre otras habilidades (Calero et al., 2018), sugiriendo que

podría haber beneficios específicos para les tutores. Sin embargo, esta cuestión no ha sido suficientemente explorada.

Modulación por teoría de la mente

Por un lado, en un sentido amplio, el aprendizaje depende del grado de desarrollo de la teoría de la mente (Carlson et al., 2013; Sabbagh & Baldwin, 2001). Por tanto, cabe esperar que la capacidad de le aprendiz para inferir el estado mental de su tutore influya sobre su experiencia de aprendizaje en la tutoría entre pares también.

Por otra parte, como resumimos anteriormente, ciertas definiciones mentalistas de la enseñanza hacen hincapié en la importancia de la teoría de la mente para enseñar, en particular para identificar brechas de conocimiento con le aprendiz. En este sentido, diferencias en esta habilidad cognitiva podrían resultar en diferencias en la capacidad de enseñar (Kline, 2014), y por consiguiente en la efectividad de la tutoría entre pares.

1.3.3.3 Metacognición

Entendemos por metacognición a la habilidad de monitorear y controlar el propio conocimiento; es decir, de reflexionar sobre el propio pensamiento, de expresar juicios de confianza sobre lo que une sabe, y de elaborar estrategias para aprender lo que une no sabe (Kornell et al., 2007; Nelson & Narens, 1990).

Efectos sobre metacognición

Por un lado, el diálogo entre tutores y aprendices promovería procesos metacognitivos (De Backer et al., 2012; King, 1998; Roscoe, 2014), que podrían resultar en mejor conocimiento acerca de cómo y cuándo usar ciertas estrategias de aprendizaje y de resolución de problemas (Shamir et al., 2008), y mejor aprendizaje auto-regulado (King, 1998; Shamir & Tzuriel, 2004). Por ejemplo, se ha reportado que una forma de tutoría entre pares, la tutoría entre pares *recíproca* (en la que se alternan los roles de aprendiz y tutore), incrementa la aplicación de estrategias de regulación metacognitiva en estudiantes universitarios, como la capacidad de estimar la calidad percibida de sus actividades (De Backer et al., 2012), y que la discusión entre elles mejora la confianza sobre sus propias respuestas (Smith et al., 2009).

Por otra parte, les aprendices podrían incorporar procesos de pensamiento y habilidades metacognitivas modelados por sus pares tutores más fácilmente de lo que ocurriría con otro tipo de instructore, debido a la proximidad entre ambes (Calero et al., 2018), como fue mencionado anteriormente.

Por último, como fue comentado para la teoría de la mente, la enseñanza ha sido postulada como motor del desarrollo metacognitivo también (Calero et al., 2018), sugiriendo beneficios específicos para les tutores. Por ejemplo, les niñes serían más proclives a reconocer que no saben cuando enseñan que cuando se les pide justificar sus respuestas (Calero, 2020).

Modulación por metacognición

Por un lado, habilidades metacognitivas insuficientes de parte de les aprendices podría limitar algunas oportunidades de aprendizaje. Por ejemplo, podrían tener dificultades para aplicar estrategias (como repreguntar a le tutore) que les permitan corregir errores de comprensión, ya sea por un control metacognitivo inadecuado, o por sobrestimar lo que saben realmente ("ilusión de conocimiento") producto de un monitoreo metacognitivo deficiente (Roscoe & Chi, 2008).

Por otra parte, las habilidades metacognitivas serían de importancia durante la enseñanza también (Calero et al., 2018; Roscoe & Chi, 2008). Problemas con el monitoreo metacognitivo podrían causar que les tutores pasen por alto errores o contradicciones en sus explicaciones, en especial si estos problemas les llevan a sobrestimar lo que saben realmente (Roscoe & Chi, 2008). Además, consecuencia de estas omisiones o producto de problemas con el control metacognitivo, les tutores podrían no responder adecuadamente a estos errores o contradicciones, limitando las oportunidades de aprendizaje de elles mismes y de sus aprendices (Roscoe & Chi, 2008).

1.3.3.4 Aspectos socioafectivos

Nos referiremos aquí a aspectos y capacidades sociales, comunicacionales, afectivas o emocionales, y motivacionales.

Efectos sobre aspectos socioafectivos

La tutoría entre pares, como forma de interacción social, involucraría y pondría en práctica aspectos y capacidades socioafectivos. Concretamente, hay evidencia de que provee ganancias en habilidades sociales y de comunicación transferibles (Xu et al., 2008), como así también en aspectos afectivos y motivacionales como el autoestima (D. Miller et al., 2010), la actitud hacia la escuela y hacia el tema en estudio (Mohan, 1972), y la satisfacción académica (K. J. Topping, 2020), tanto para aprendices como para tutores.

Modulación por aspectos socioafectivos

Recíprocamente, cabe esperar que la tutoría entre pares a su vez dependa de los aspectos y capacidades socioafectivos que involucra y pone en práctica. Por ejemplo, la empatía podría ser una habilidad importante para les tutores (J. Cohen, 1986). Además, como comentamos anteriormente, una relación de amistad entre tutore y aprendiz podría promover discusiones críticas y productivas en tareas de razonamiento científico (Azmitia & Montgomery, 1993), y la confianza que le aprendiz tiene en su tutore sería parte importante del éxito de le tutore (Corriveau & Winters, 2019).

1.3.4 Formas de implementación

Como hemos visto, la tutoría entre pares ha sido asociada con numerosos beneficios. Sin embargo, simplemente juntar estudiantes y esperar lo mejor, aunque pueda tener efectos sutiles, no es lo óptimo (K. J. Topping, 2020). Cuando decimos "*tutoría entre pares"*, en realidad nos estamos refiriendo a un conjunto bastante heterogéneo de prácticas o implementaciones concretas (Kalkowski, 1995), caracterizadas por una larga serie de variables o dimensiones (Bowman-Perrott et al., 2013; Leung, 2015; Roscoe & Chi, 2007; K. J. Topping, 2005). Esto complejiza el estudio de este tipo de enseñanza y hace que los resultados sean altamente variables, y a veces hasta contradictorios.

En la sección anterior consideramos parámetros de les participantes o *variables de contexto*, individuales o de grupo, que podrían ser afectadas por y moduladoras de la tutoría entre pares, en una suerte de ciclo virtuoso. Aquí mencionaremos parámetros de la metodología o *variables de organización* (Bowman-Perrott et al.,

2013; Leung, 2015), características de cada implementación concreta, que por su naturaleza estructural no caben ser afectadas por la práctica misma, pero que pueden modificar su desarrollo:

- **Dominio del conocimiento:** La tutoría entre pares ha demostrado ser efectiva para el aprendizaje en varios dominios del conocimiento, incluyendo lectura, matemática, ciencias sociales, ciencias naturales, entre otros (Bowman-Perrott et al., 2013; Leung, 2015; K. Topping, 2017).
- **Edad y nivel educativo:** Asimismo, ha demostrado ser útil para diversas edades y niveles educativos, desde el jardín de infantes hasta la universidad, aunque los efectos más grandes habitualmente se han encontrado en la escuela secundaria (Leung, 2015).
- **Género:** La combinación de género de les participantes sería una variable de importancia. Concretamente, parejas con tutore y aprendiz del mismo género estarían asociadas con efectos significativamente mayores que parejas *mixtas*, encontrándose dificultades particularmente en parejas con aprendices de género masculino y tutoras de género femenino (Leung, 2019).
- **Entrenamiento para tutores:** La disponibilidad de capacitaciones para les tutores no siempre implica tutorías entre pares más efectivas, sino que depende también de la frecuencia y duración de estas capacitaciones (Leung, 2015, 2019).
- **Estructura:** Entendemos por tutoría entre pares estructurada aquella que dispone de guiones, materiales estructurados, y/o procedimientos claros para la interacción (K. J. Topping, 2020). En general, este tipo de implementaciones están asociadas a mejores resultados (P. A. Cohen et al., 1982; Leung, 2015, 2019).

La lista de variables continúa, habiéndose considerado también otros aspectos como el contexto socioeconómico de les participantes, el número, frecuencia y duración de las instancias de interacción, la presencia y el tipo de recompensas, la reciprocidad en la asignación de roles, entre otros (Bowman-Perrott et al., 2013; P. A. Cohen et al., 1982; Leung, 2015, 2019).

Dada la complejidad del escenario planteado, la descripción adecuada de estas variables en los estudios sobre tutoría entre pares es fundamental para su evaluación y comparación, y para la identificación de factores relevantes (K. J. Topping, 2020).

1.3.5 Limitaciones

Como se desarrolló en las secciones anteriores, la tutoría entre pares es un tema ampliamente estudiado. Sin embargo, aún presenta numerosos aspectos sin resolver.

Por un lado, ha demostrado efectos sobre el aprendizaje, tanto para aprendices como para tutores, en diversos dominios del conocimiento y niveles educativos. Sin embargo, la mayoría de estos estudios se han enfocado en matemática y lectura (Leung, 2015), y algunos dominios, de relevancia para el contexto sociocultural actual, permanecen subexplorados o sin explorar, en particular para algunas edades y niveles educativos. Además, si bien se han reportado diversos efectos sobre les tutores, no existe un consenso claro con respecto a las condiciones bajo las cuales éstos se hacen evidentes, algunos estudios no han podido encontrarlos (de la Hera et al., 2019; Rohrbeck et al., 2003; Roscoe & Chi, 2007), y hay quienes incluso temen que podrían existir perjuicios para les tutores (Kalkowski, 1995). Algo similar ocurre con el impacto a largo plazo. Es decir, a pesar de haberse observado efectos, en particular de retención de niveles de conocimiento, no habrían sido lo suficientemente estudiados en la literatura sobre aprendizaje colaborativo y enseñanza entre pares (K. J. Topping, 2020).

Por otra parte, en las secciones precedentes hemos comentado acerca del efecto sobre y la modulación por otros aspectos cognitivos y socioafectivos. En particular, cabe destacar la supuesta relación entre tutoría entre pares y metacognición. Si bien esta relación ha sido relativamente respaldada por algunos estudios, aún quedan aspectos por explorar, como la posibilidad de que la enseñanza motorice su desarrollo (Calero et al., 2018).

Finalmente, muchas investigaciones fallan en describir en detalle la implementación específica empleada. Dada la complejidad en torno a qué entendemos por tutoría entre pares, esto complica la comparación entre estudios. Si queremos tener un

mejor entendimiento acerca de cómo funciona la tutoría entre pares, necesitamos estudios mejor definidos, que describan adecuadamente el tipo de tutoría entre pares empleado, el contexto y la muestra (K. J. Topping, 2020).

A continuación presentamos los objetivos del presente trabajo, en contexto del marco teórico establecido y sus limitaciones.

1.4 Estudios experimentales de la enseñanza entre pares en la niñez

En este trabajo indagaremos acerca de cómo podemos optimizar ciertos aspectos del intercambio y la construcción de conocimiento, contribuyendo así al progreso de la pedagogía y la participación cultural. Con este fin, ofreceremos evidencia empírica para el uso de la enseñanza entre pares como herramienta en casos concretos de relevancia pedagógica.

1.4.1 Objetivos específicos

Específicamente, los objetivos del trabajo son realizar estudios experimentales controlados que permitan:

1. Evaluar el aprendizaje asistido por pares en general, y la tutoría entre pares en particular, en dominios del conocimiento relevantes para el contexto sociocultural actual en los que no ha sido explorada, o en los que ha sido explorada insuficientemente.
2. Identificar diferencias entre aprender de une par y aprender de une adulte, que permitan subrayar el rol complementario de la tutoría entre pares y otras estrategias de enseñanza tradicional.
3. Evaluar efectos sobre aprendices y tutores, a corto y largo plazo.
4. Explorar los efectos sobre y la modulación por habilidades metacognitivas.
5. Describir adecuadamente y tener en cuenta variables de contexto, específicamente edad y género, y de organización, concretamente la brecha

de conocimiento entre tutores y aprendices, y el tipo específico de intervención.

1.4.2 Hipótesis de trabajo

En función de los objetivos específicos, planteamos las siguientes hipótesis de trabajo, siempre en contexto de las condiciones particulares de los estudios desarrollados:

1. La tutoría entre pares reduce la brecha entre participantes con distintos niveles de conocimiento.
2. Sin perjuicio de esta reducción, los aprendizajes derivados de la tutoría entre pares presentan diferencias respecto de los derivados de formas más tradicionales de enseñanza, confirmando su rol complementario.
3. Se observan efectos tanto en aprendices como en tutores, y tanto a corto plazo como a largo plazo.
4. Es posible identificar una relación bidireccional entre la tutoría entre pares y algunas habilidades metacognitivas.
5. Los efectos descriptos se observan para distintas edades y géneros, pero dependen de la amplitud de la brecha de conocimiento entre participantes.

1.4.3 Desarrollo

Con el fin de alcanzar los objetivos descriptos realizamos tres estudios experimentales. Todos ellos son ensayos controlados, aleatorizados, con variables de contexto y de organización bien definidas, diseñados para modelar interacciones ecológicamente válidas (Bronfenbrenner, 1979).

En líneas generales, los estudios comprenden (1) una primera etapa de evaluación para estimar el nivel de conocimiento inicial, seguida de (2) una etapa de interacción (o control sin interacción), y (3) una etapa de evaluación final. En todos los casos el control sin interacción fue un control activo, con una tarea de revisión análoga pero individual. Además, en ocasiones se incluyó una etapa previa de instrucción para una parte de les participantes, y/o etapas adicionales para evaluar variables potencialmente relacionadas. Si bien todas las actividades de interacción fueron

estructuradas (ver sección “1.3.4 Formas de implementación”), en ningún caso les participantes recibieron formación como tutores.

Nuestros estudios involucran más de 200 niñes de entre 6 y 12 años de edad, de escuelas primarias del Área Metropolitana de Buenos Aires. El rango de edad seleccionado busca poder atender las exigencias de los procedimientos, y al mismo tiempo satisfacer los objetivos propuestos en relación a la consideración de distintas edades. Los estudios difieren principalmente en el dominio o área de conocimiento, y en algunos casos en la forma específica de la actividad de intercambio y colaboración.

Puntualmente, en el primer estudio (Capítulo 2) nos enfocamos en el aprendizaje asistido por pares como medio para facilitar procesos de cambio conceptual, a menudo difíciles de lograr, de importancia en el desarrollo de les niñes, y en el contexto de crisis de posverdad, desinformación y negacionismo científico del contexto sociocultural en que vivimos. ¿Puede la enseñanza entre pares ayudar a promover estos procesos?

En el segundo estudio (Capítulo 3) nos enfocamos de lleno en la tutoría entre pares propiamente dicha, como herramienta para promover el aprendizaje en un dominio de relevancia para el mundo actual: la programación de computadoras. ¿Puede la enseñanza entre pares promover el aprendizaje de estas competencias, clave para el mundo de hoy?

Finalmente, en el tercer estudio (Capítulo 4), restamos importancia al dominio y empleamos uno caracterizado por su simplicidad de análisis, permitiendo así la consideración de aún más factores potencialmente relacionados con la tutoría entre pares. ¿Podemos optimizar el estudio experimental de la enseñanza entre pares, ampliando la muestra y el número de variables consideradas?

Desarrollamos a continuación estos tres estudios principales, uno por capítulo, acompañados además por estudios menores que los complementan.

Capítulo 2: Enseñanza entre pares y cambio conceptual

A lo largo de su desarrollo, les niñes atraviesan procesos de transición conceptual, a menudo reemplazando conocimiento intuitivo por teorías científicamente aceptadas. Con frecuencia esto genera situaciones de conflicto social, en las que distintes niñes pueden sostener teorías y explicaciones significativamente diferentes acerca de un mismo fenómeno.

En este capítulo exploraremos si la enseñanza entre pares y el intercambio de argumentos pueden ser catalizadores del cambio conceptual. En particular, sobre los modelos mentales que les niñes tienen sobre el planeta Tierra.

2.1 Introducción

2.1.1 Cambio conceptual

2.1.1.1 El origen de los conceptos

Desde la infancia, les niñes tienen una comprensión intuitiva de cómo funciona el mundo, incluyendo nociones elementales de física, matemática y entidades sociales (Carey, 2009; Gopnik et al., 1997). La mayoría de les investigadores coinciden en que estas intuiciones generales, en parte construidas sobre la base de observaciones y experiencias cotidianas (Carey & Spelke, 1994; Keil, 1994; Piaget, 1929), forman conjuntos coherentes y sistemáticos de ideas. Estos conjuntos corresponden a dominios bien definidos del pensamiento humano (física, psicología, matemática, lenguaje, etc), cada uno con su propia ontología, y se conocen como *teorías marco*. Estas teorías, en principio ingenuas, a menudo difieren de las aceptadas científicamente y –como veremos a continuación– son resistentes al cambio (Hadzigeorgiou, 2015; Saçkes, 2015; Vosniadou, 2009). Además, condicionan la interpretación de nuevas observaciones y la elaboración de teorías más específicas, y son la base de modelos mentales que les niñes usan para dar explicaciones causales y hacer predicciones acerca del mundo (ver Figura 2.1).

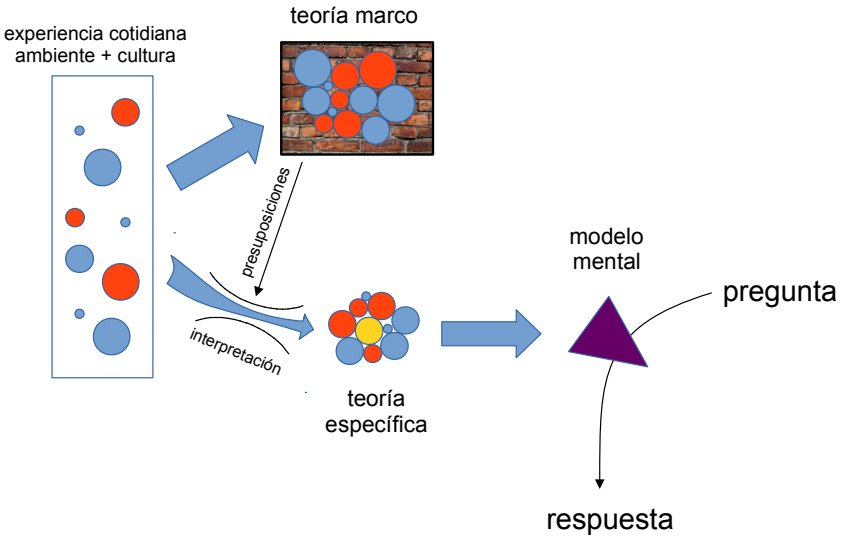


Figura 2.1: *Perspectiva del conocimiento como teoría*

Les niñes organizan la información obtenida de su entorno (círculos rojos y azules) en conjuntos coherentes o *teorías marco*. Las presuposiciones de estas teorías a su vez limitan la interpretación de nuevas observaciones (círculo amarillo) y la construcción de *teorías específicas*. Es a partir de estas teorías que les niñes elaboran y manipulan *modelos mentales* sobre el mundo, para responder preguntas y realizar predicciones. Según esta perspectiva del *conocimiento como teoría*, el cambio conceptual ocurre por enriquecimiento y/o revisión de las teorías elaboradas.

Este enfoque, del conocimiento organizado en conjuntos coherentes o teorías, es la perspectiva del *conocimiento como teoría* (Blown, 2013; Carey, 2009; Eikospentaki et al., 2015; Vosniadou et al., 2009), que adoptaremos en este trabajo. Sin embargo, cabe mencionar que existen otras perspectivas, minoritarias, como la del *conocimiento como fragmentos* (diSessa, 2009; Saçkes, 2015), según la cual les niñes no tienen concepciones previas, acertadas o erradas, sino que están libres de teorías, con su conocimiento fragmentado en vez de coherente (diSessa, 2009; Frède et al., 2011; Panagiotaki et al., 2009).

2.1.1.2 Procesos de cambio conceptual

Los procesos de *cambio conceptual* ocurren en contextos de aprendizaje y de desarrollo del conocimiento (Thagard, 2009; Vosniadou et al., 2009). Abarcan fenómenos individuales y sociales, y algunos involucran transiciones conceptuales

radicales en las que modelos conceptuales arraigados colapsan y son reemplazados por otros (Carey, 2009; T. S. Kuhn, 1971). En particular, el acercamiento de Teoría Marco al cambio conceptual (del inglés, *Framework Theory approach to conceptual change*), lo define como cualquier cambio de conocimiento que ocurre en procesos de aprendizaje y desarrollo, desde el simple enriquecimiento de las estructuras conceptuales preexistentes, hasta la compleja revisión radical de las teorías marco (Vosniadou, 1994; Vosniadou et al., 2009).

Es decir, a lo largo de su desarrollo (y durante toda su vida), y como producto de la interacción con su entorno natural y cultural, les niñes enriquecen sus teorías o, frente a observaciones contradictorias, las revisan y abandonan para elaborar otras nuevas (Carey, 1991, 2009; Gopnik et al., 1997; Hadzigeorgiou, 2015; Saçkes, 2015; Vosniadou et al., 2009).

2.1.1.3 Cambio conceptual en astronomía observacional

En este trabajo nos enfocamos en el dominio de la astronomía observacional, la cosmología y las miradas sobre el mundo (referido de aquí en más simplemente como *dominio de la astronomía observacional*), un campo que en los últimos años ha visto un rápido crecimiento del pensamiento pseudocientífico, con movimientos internacionales volviendo a instalar ideas hace siglos obsoletas (Landrum, 2019; “The Challenge of the Post-Truth Era”, 2018). Les niñes poseen una amplia diversidad de cosmologías y modelos mentales de la Tierra, incluyendo su forma, dónde vivimos en ella, y cómo ésta se relaciona con el espacio que la rodea (Mali & Howe, 1979; J. Nussbaum, 1979; Sneider & Pulos, 1983). Además, este dominio también es interesante por su elaborada red de conceptos y relaciones conceptuales, y porque el campo en sí mismo ha sufrido restructuraciones radicales a lo largo de su desarrollo histórico, y es posible que los cambios experimentados por les niñes sean análogos a aquellas (Hadzigeorgiou, 2015; Vosniadou & Brewer, 1992).

El proceso de cambio conceptual, desde un modelo mental inicial de planeta plano al de una comprensión acabada de planeta esférico, ocurre a través de la revisión de las presuposiciones de las teorías marco. Éste es un proceso lento y gradual, durante el cual les niñes intentan acomodar en sus estructuras conceptuales preexistentes la información que reciben de sus experiencias cotidianas, dando

origen a modelos mentales alternativos inexactos condicionados por sus presuposiciones (Vosniadou, 1994) (ver ejemplos en Figura 2.2). Los modelos más simples se dan a edades más tempranas, mientras que los más sofisticados se dan en niñes mayores (Vosniadou et al., 2009). En suma, el proceso de cambio conceptual involucra un cambio de categorización ontológica de la Tierra, para reconceptualizarla de un objeto físico (al que aplican las mismas presuposiciones de la teoría marco de la física que aplican a otros objetos físicos en la Tierra) a un objeto astronómico (Vosniadou et al., 2005, 2009). Este proceso es análogo a como las teorías científicas cambian a lo largo del tiempo, y la astronomía es precisamente un claro ejemplo de ello (T. S. Kuhn, 1957).

Aunque los caminos de aprendizaje de les niñes son complejos, y el desarrollo de su comprensión conceptual no es lineal o jerárquica (Saçkes, 2015), es posible identificar una ruta más o menos general, expuesta sucintamente a continuación:

1. Al principio del camino de cambio conceptual el concepto de la Tierra que tienen les niñes está condicionado por las presuposiciones de sus teorías marco sobre la física. Por ejemplo, suponen que las personas viven sobre suelo plano, y que los objetos sin sostén caen hacia abajo, noción que adquieren alrededor de los cuatro meses de edad (Spelke et al., 1992). Estas presuposiciones limitan la interpretación de sus observaciones y de la información cultural que reciben (Hadzigeorgiou, 2015), condicionando así la elaboración de *teorías específicas* sobre la Tierra (Vosniadou, 1994; Vosniadou et al., 2009) y los modelos mentales correspondientes.
2. A continuación, mientras continúan acomodando la información que reciben del entorno, suelen elaborar modelos de *Tierra dual*, en los que asumen que les adultes se refieren a una entidad diferente cuando hablan de la “Tierra redonda en el espacio”. Éstos no requieren la revisión de presuposiciones, sino la simple incorporación de información cultural (científica) novedosa a sus estructuras conceptuales preexistentes.
3. Luego, más adelante en el desarrollo, revisan su presuposición sobre la caída de objetos sin sostén, consiguiendo suspenderla en cuanto a la Tierra respecta, y permitiéndose entonces conceptualizarla como un objeto suspendido en el espacio. Sin embargo, esta presuposición sigue en pie para

los objetos (incluyendo a las personas) en la superficie terrestre, condicionando el tipo de modelos de *Tierra suspendida* posibles, y dando lugar a modelos de disco plano, o de esfera hueca con las personas viviendo en suelo plano *dentro* de la esfera.

4. Finalmente, tras reinterpretar las presuposiciones de caída hacia abajo y de suelo plano, logran conceptualizar a la Tierra, dentro del marco explicativo de los objetos solares/astronómicos, como una esfera con las personas viviendo en toda su superficie, sostenidas por su atracción gravitatoria (Blown, 2013; Vosniadou & Brewer, 1992).

2.1.1.4 Importancia pedagógica del cambio conceptual

Aprender en ciencia, entonces, implica procesos de cambio conceptual. Sin embargo, estos procesos pueden ser difíciles de lograr (Damon & Phelps, 1989), especialmente cuando implican la revisión de teorías marco (Vosniadou, 1994), resaltando la necesidad de estrategias de enseñanza específicas (Hadzigeorgiou, 2015; Hestenes et al., 1992).

Estos procesos graduales, desde nociones intuitivas hasta conocimientos más avanzados, implican acomodar información nueva dentro de estructuras conceptuales preexistentes, dando lugar a modelos mentales sintéticos alternativos condicionados por las presuposiciones de les niñes (Vosniadou & Skopeliti, 2014). Como resultado, en contextos educativos es común encontrar niñes en diferentes puntos a lo largo de estos caminos del desarrollo (Vosniadou & Brewer, 1992). Esto conduce a situaciones de disparidad del conocimiento, en las que distintes niñes poseen teorías sustancialmente diferentes acerca de un mismo fenómeno.

En este trabajo evaluamos cómo estas diferencias podrían ser aprovechadas para promover el desarrollo de estos difíciles procesos.

2.1.2 Interacción social y construcción de conocimiento

En esta sección se discutirá la interacción social como motor de la construcción de conocimiento, tanto colectivo como individual. En particular, hablaremos del *conflicto sociocognitivo* como una forma de interacción social, y de cómo no todas las formas de resolverlo son productivas para la construcción de conocimiento. Se mencionarán

además posibles aspectos negativos de la interacción social, de especial importancia en contexto de su utilización como herramienta pedagógica.

2.1.2.1 Conflicto sociocognitivo

Las personas a menudo tenemos conocimiento y opiniones diferentes y contrastantes. La ciencia no es una excepción, y habitualmente una variedad de teorías e interpretaciones existen para un mismo asunto. La forma en la que grupos de personas resuelven los conflictos y desacuerdos que resultan de estas diferencias depende de la naturaleza de la situación, implicando diversas estrategias que conducen a distintos resultados en la construcción colectiva de conocimiento.

Además, el conflicto no sólo contribuye a la construcción colaborativa de conocimiento, sino que es sabido que también promueve el desarrollo cognitivo individual (Heyd-Metzuyanim & Schwarz, 2017; Piaget, 1985). La cuestión de la construcción de conocimiento individual ha sido abordada tanto por el *cognitivismo* Piagetiano como por el *socioconstructivismo* Vygotskiano, y sus derivaciones (Heyd-Metzuyanim & Schwarz, 2017). Aunque ambos comparten una mirada *constructivista*, con les sujetos construyendo conocimiento activamente, hay diferencias entre ellos. El primero se enfoca en el conflicto (socio)cognitivo y en el despliegue de tendencias naturales, que resultan de la interacción con el mundo externo, incluyendo el mundo social (S. A. Miller & Brownell, 1975; Mugny & Doise, 1978). El segundo, en cambio, se enfoca en el mundo externo en sí mismo, representado por la cultura circundante, prestando especial atención a las diferencias de competencias con le sujeto (Castellaro & Roselli, 2018; Heyd-Metzuyanim & Schwarz, 2017).

Cuando las personas en un grupo tienen opiniones diferentes y tratan de llegar a un consenso, usan distintas estrategias con distintos resultados (Butera et al., 2011), como desarrollaremos a continuación.

2.1.2.2 Formas de resolución del conflicto

Una posibilidad es elegir una combinación promedio entre todas las opciones. Es bien conocido que, dependiendo de la naturaleza de la situación, elegir la respuesta promedio de un grupo suficientemente grande de respuestas independientes es mejor que escoger cualquiera de las respuestas individuales, un fenómeno

popularmente conocido como sabiduría de las masas (Galton, 1907). Sin embargo, habitualmente las respuestas no son independientes (por ejemplo, cuando las personas intercambian información) y en estos casos el efecto de sabiduría de las masas se rompe (Lorenz et al., 2011).

Por otra parte, en algunos casos la interacción entre les miembros de un grupo –en vez de simplemente observar el comportamiento de les otres– puede fortalecer el desempeño del grupo por encima del de sus partes. Por ejemplo, en tareas de decisión perceptual de bajo nivel, grupos de dos personas que interactúan libremente para llegar a un consenso son mejores que les individues aislades, posiblemente porque comparten sus niveles de confianza acerca de sus respuestas (Bahrami et al., 2010). Un efecto de deliberación positivo similar, probablemente mediado por heurísticas o suposiciones de confianza y habilidad, fue encontrado en la estimación de magnitudes de conocimiento general, tales como la altura de la Torre Eiffel, tanto en las respuestas grupales, como en las individuales después de la interacción (Navajas et al., 2018).

Sin embargo, cuando las decisiones grupales se basan en el razonamiento, la comunicación no se enfoca en la confianza individual o en la supuesta pericia, sino en el intercambio de argumentos (Mercier & Sperber, 2012). El debate y la deliberación en grupos pequeños de adultes resultan en un marcado aumento en la calidad de las decisiones que toman estos grupos (Navajas et al., 2018; Shi et al., 2019), y este procedimiento parece apropiado para grupos de niñes, que también son capaces de producir, compartir y evaluar argumentos (Mercier, 2011). Es a través de este tipo “epistémico” de elaboración del conflicto, basado en argumentaciones y justificaciones, que se obtendrían resultados cognitivos positivos, en contraste con los tipos “relacionales” discutidos arriba, que se enfocan en la confianza o experiencia de les individues (Buchs et al., 2004; Castellaro & Roselli, 2018; Mercier, 2011).

En particular, las diferencias en el conocimiento científico son especialmente abordadas mediante decisiones grupales basadas en el razonamiento y la argumentación, y existe abundante evidencia a favor de que argumentar sirve un rol en el aprendizaje de conceptos científicos esquivos (Osborne, 2010). Mientras construyen conocimiento, les niñes pueden involucrarse en argumentación colaborativa, en la que trabajan juntes para construir y criticar argumentos, mientras

permanecen libres para explorar posiciones con flexibilidad y hacer concesiones (Asterhan & Schwarz, 2009; E. M. Nussbaum, 2008). Esto puede promover cambios conceptuales y del desarrollo en niñes (D. Kuhn, 2015; Mugny & Doise, 1978; E. M. Nussbaum, 2008), en parte debido a las demandas cognitivas de elaboración que supone la argumentación (Coleman et al., 1997; E. M. Nussbaum, 2008). De hecho, las investigaciones parecen sugerir que para inducir cambios conceptuales a través de la colaboración entre pares, es necesario que exista argumentación entre elles (Asterhan & Schwarz, 2009).

2.1.2.3 Efectos negativos de la interacción social

Sin embargo, la evidencia también indica que las interacciones sociales pueden conducir a comportamientos colectivos irracionales. Por ejemplo, pueden provocar cascadas de información que lleven a comportamiento gregario (Mackay, 1841; Raafat et al., 2009), o potenciar la difusión masiva de noticias falsas (Lazer et al., 2018) o ideas pseudocientíficas (Kata, 2010), opacando así sus efectos benéficos.

Por lo tanto, la cuestión de si la interacción entre niñes con miradas diferentes les hace progresar hacia modelos basados en la razón y en evidencia científica es una pregunta abierta que requiere análisis empírico. Éste es el objetivo del estudio principal de este capítulo, en el que exploramos el potencial de la enseñanza entre pares como medio para elaborar el conflicto sociocognitivo y promover procesos de cambio conceptual.

2.2 Estudio principal 1

En este estudio evaluaremos la enseñanza entre pares como estrategia pedagógica para promover el cambio conceptual en astronomía observacional[3]

2.2.1 Introducción

Cuando les miembros de un grupo tienen opiniones contrastantes sobre un asunto, el conflicto sociocognitivo que sobreviene podría conducir a procesos de desarrollo

3 Los resultados de este estudio fueron publicados en: de la Hera, D. P., Sigman, M., & Calero, C. I. (2019). Social interaction and conceptual change pave the way away from children's misconceptions about the Earth. *npj Science of Learning*, *4*(1). https://doi.org/10.1038/s41539-019-0051-3

del conocimiento, central para que ocurran fenómenos de cambios conceptual (C. Howe et al., 1990; Limón, 2001; E. M. Nussbaum, 2008). Si bien existe evidencia de la colaboración como herramienta para el desarrollo intelectual (Curtis & Lawson, 2019; Laal & Laal, 2012), hay pocos estudios experimentales rigurosos que comparen grupos *versus* individues participando de una tarea similar, y que demuestren ventajas para la condición grupal. Especialmente en el caso de tareas poco estructuradas, que pueden no tener una única respuesta correcta, como la que presentaremos a continuación (D. Kuhn, 2015).

El objetivo de este trabajo es comprender el rol de la enseñanza entre pares, como caso particular de interacción social y conflicto sociocognitivo, como estrategia pedagógica para promover la construcción colectiva de conocimiento, y los procesos individuales de cambio conceptual, cuando niñes con distintos modelos sostienen y comparten argumentos a favor de estos modelos. En particular, nos enfocaremos en el dominio de la astronomía observacional, aprovechando la diversidad de modelos conceptuales del planeta que existen naturalmente dentro de un aula.

Con este fin, evaluamos y cuantificamos los modelos mentales de la Tierra de un grupo de niñes antes y después de su interacción entre pares, y los comparamos con los de un grupo control, en el que les niñes realizaron la misma tarea a solas, sin interacción. Antes de esta interacción, en una primera entrevista, se evaluó el conocimiento de cada niñe a través de preguntas tanto fácticas (referidas a hechos concretos) como generativas, y luego fueron agrupados en parejas con une participante menos experimentade, grupo *Inexperto* (I), y otre más experimentade, grupo *Avanzado* (A). La interacción entre pares consistió en un trabajo colaborativo en el que se pidió a les niñes que expresaran consenso –a través de un dibujo– en relación a varios aspectos geográficos y astronómicos usados para explorar su nivel de comprensión sobre nuestro planeta.

Debido a la naturaleza del dominio de conocimiento seleccionado, todes les participantes tenían algún conocimiento previo. Por lo tanto, para favorecer flujos de información similares a los esperables en tutorías entre pares tradicionales, se incorporaron asimetrías de roles explícitas. De este modo, la tarea no sólo respondía al paradigma neopiagetiano de conflicto sociocognitivo (Buchs et al., 2004; Limón, 2001; Mugny & Doise, 1978), sino que además incluía asimetrías neovygotskianas de roles y de niveles de conocimiento, en línea con propuestas previas para

reconciliar las miradas cognitivista y socioculturalista del cambio conceptual (Heyd-Metzuyanim & Schwarz, 2017).

El diseño experimental pretende cuantificar: (a) si les niñes cambian su comprensión conceptual de la Tierra después de un trabajo colaborativo; (b) si las representaciones mentales de ambes niñes en cada pareja se hacen más similares después de la interacción; y (c) si este cambio refleja una evolución hacia representaciones más avanzadas del conocimiento, en sintonía con nuestra comprensión física y cosmológica actual.

2.2.2 Métodos

2.2.2.1 Participantes

Un total de 46 estudiantes de dos clases de segundo grado (23 de género masculino, y 23 femenino, según fue informado por las autoridades escolares) de una escuela bilingüe en Buenos Aires de nivel socieconómico medio-alto y alto participaron del estudio. Los datos de cuatro participantes (3 niños y 1 niña) tuvieron que ser descartados debido a problemas técnicos. Las edades de les 42 participantes restantes comprendieron entre 7 años y 8 años y 3 meses, con una edad media de 7 años y 6 meses. Las sesiones se llevaron a cabo fuera del aula, en una sala controlada provista por la escuela para uso exclusivo de este estudio.

Les padres o responsables legales de todes les participantes firmaron un consentimiento voluntario previamente autorizado por un comité de ética: Comité de Ética de la Dirección de Investigación del Centro de Educación Médica e Investigación Clínica “Norberto Quirno” (CEMIC), Unidad Asociada del CONICET, Protocolo N° 683.

2.2.2.2 Software y equipamiento

Todas las sesiones fueron registradas con dos cámaras web *Logitech HD Pro C920* con micrófonos integrados. Los dibujos fueron realizados en una pantalla *Wacom* con un lápiz electrónico sensible a la presión y a la inclinación. El lienzo de dibujo fue generado con el software *GIMP*, y video, audio y pantalla fueron simultáneamente registrados con el software *VLC* de *VideoLan*, en una computadora con *Linux Mint 17*.

2.2.2.3 Diseño experimental

Les niñes fueron entrevistades dos veces para evaluar sus modelos mentales de la Tierra antes y después de la interacción entre pares. Primero se usó un esquema de codificación general para evaluar a les participantes rápidamente después de la Entrevista 1 (ver a continuación) y agruparlos en parejas de niñes con distintos niveles de conocimiento. Un segundo esquema de codificación detallado se usó más adelante para evaluar cambios de conocimiento entre las entrevistas.

Entonces, cada niñe participó de tres sesiones: (a) Entrevista 1 (*pre-test*), (b) interacción entre pares o revisión autónoma, y (c) Entrevista 2 (*post-test*).

Entrevista 1

Todes les participantes fueron entrevistades individualmente por une de dos investigadores, siguiendo un cuestionario estrictamente guionado. El diseño del cuestionario se basó en el trabajo de Vosniadou y Brewer (1992). Incluía cincuenta preguntas para revelar los modelos mentales de la Tierra de les participantes (ver Anexo 2.1: Cuestionario completo). Las preguntas exploraban temas tales como la forma y la ubicación de entidades astronómicas (planeta Tierra, luna, estrellas, sol), del cielo (cielo, nubes) y de la Tierra (personas, casas, árboles, países, mar), como así también la exploración de la superficie terrestre (marcha en línea recta y “bordes” del planeta). Como es común que les niñes cambien sus respuestas si se les pregunta repetidas veces (Cassel et al., 1996; Siegal, 1991), la estrategia precisa para preguntas adicionales y de seguimiento fue predefinida para minimizar sesgos introducidos por le entrevistadore. Les participantes fueron familiarizades con la pantalla de dibujo antes de comenzar con la Entrevista 1, y eran libres de dibujar, hablar y gesticular para responder a todas las preguntas, para una mejor apreciación de su conocimiento (Calero et al., 2019; Crawford, 2005; Kelly et al., 1998; Perry et al., 1988),

Esquema de codificación 1 general

Para asegurarnos de que los modelos mentales revelados en la Entrevista 1 cambiaran lo menos posible antes de la etapa de interacción entre pares o revisión autónoma, el tiempo entre estas dos etapas debía ser corto. Para poder evaluar el conocimiento de les participantes y formar las parejas en sólo dos días, usamos un esquema de codificación general que permitió evaluarles rápidamente. El conjunto

de respuestas de cada niñe fue codificado en tres ejes: hueco/no-hueco, discoidal/esférico, y dual/no-dual, basados en los modelos mentales descriptos previamente (Vosniadou & Brewer, 1992). Se consideraron respuestas verbales, dibujadas y gestuales, y se asignó un puntaje entre -3 y +3 en cada uno de los tres ejes, según una escala tipo Likert de 7 puntos. Estos puntajes fueron asignados después de considerar el video de cada entrevista y de recolectar evidencia para decidir signo (+/-) y valor (1-3) para cada eje de forma independiente. Puntajes negativos indicaban que las respuestas de le niñe correspondían a un modelo de la Tierra hueco, discoidal, y/o dual, respectivamente. Por el contrario, puntajes positivos se usaron para participantes cuyas respuestas fueran compatibles con un modelo no hueco, esférico, y/o no dual. Un valor de cero en cualquiera de los tres ejes indicaba que las respuestas eran ambiguas, y valores cercanos a -3 o +3 reflejaban que las respuestas correspondían claramente a uno de los dos extremos opuestos de los ejes. Por ejemplo, la respuesta “hacia arriba o con un telescopio” a la pregunta P05 “¿A dónde tenemos que mirar para ver el planeta Tierra?” era tomada como evidencia a favor de un modelo dual y por lo tanto inclinaba el puntaje en el eje dual/no dual hacia -3. En cambio, responder “hacia abajo” y/o señalar hacia el piso en respuesta a la misma pregunta inclinaba el puntaje hacia +3. Ejemplos representativos de los extremos de los ejes se muestran en la Figura 2.2. Finalmente, los puntajes en cada eje fueron linealmente normalizados al rango 0-1; es decir, por ejemplo, un puntaje de -3 fue normalizado como 0, y uno de +3 como 1.

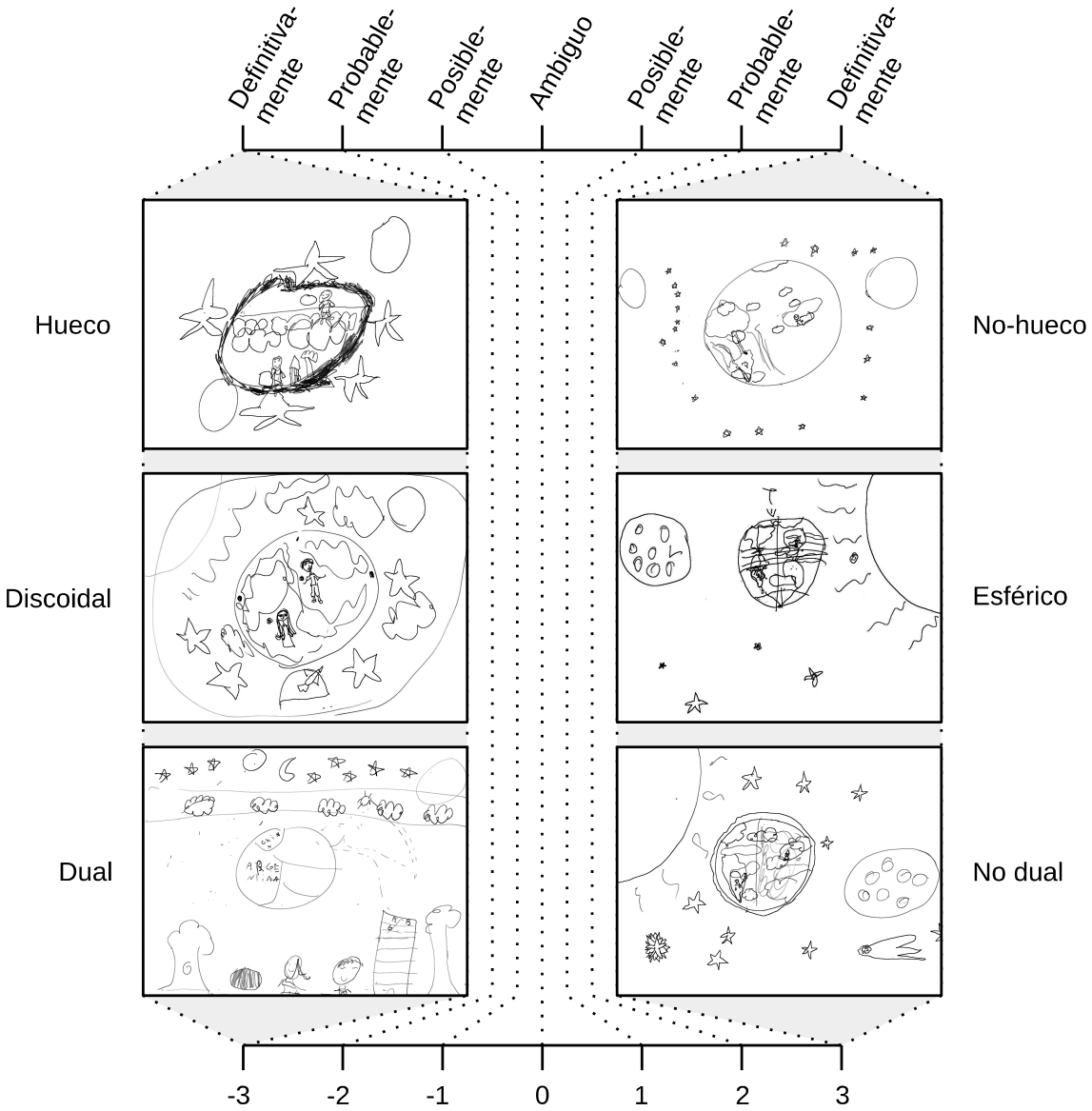


Figura 2.2: *Esquema de codificación 1, general*

Los modelos manifestados por les participantes fueron calificados primero con el Esquema de codificación 1 para agruparles en parejas para la interacción entre pares. Este esquema posee tres ejes: hueco/no hueco, discoidal/esférico, dual/no-dual; cada uno puntuado en una escala tipo Likert de 7 puntos. En la figura se muestran ejemplos representativos de modelos extremos para cada uno de los ejes.

Este esquema se basa en modelos mentales previamente descriptos en niñes: (1) la Tierra plana rectangular: con las personas viviendo sobre un suelo plano que se extiende infinitamente hacia abajo; (2) la Tierra discoidal: en la que el suelo plano tiene forma de disco; (3) la Tierra dual: con una Tierra plana en la que viven las personas y otra suspendida en el espacio; (4) la Tierra esférica hueca: con la Tierra suspendida en el espacio y las personas viviendo sobre suelo plano dentro de ella;

y (5) la Tierra esférica aplanada: con la gravedad sujetando a las personas que viven a ambos lados de una esfera aplanada (Mali & Howe, 1979; J. Nussbaum, 1979; Sneider & Pulos, 1983; Vosniadou & Brewer, 1992).

Es importante aclarar que este esquema de codificación general se usó solamente para agrupar a les participantes de a pares. De todos modos, más adelante se pudo confirmar que resultó una aproximación apropiada del segundo esquema de codificación detallado utilizado en los análisis (ver apartado *Esquema de codificación 2 detallado* de la sección “2.2.2.3 Diseño experimental”).

Usando el esquema de codificación general descripto, les participantes fueron agrupades en parejas del mismo género, tal que en cada par une de les miembros tenía una mejor comprensión (participante A, *Avanzado*) que le otre (participante I, *Inexperto*) del tema en cuestión. Es decir, le participante A tenía un puntaje más alto en al menos uno de los tres ejes del esquema de codificación, e igual o más alto en los otros dos. Para asegurarnos de que les miembros de ambas condiciones (ver a continuación) fueran elegides aleatoriamente, de estas parejas, quince fueron seleccionadas al azar para formar parte de la condición *de interacción entre pares*. Las parejas seleccionadas fueron presentadas a las docentes para confirmar que ninguna fuera especialmente conflictiva según su experiencia diaria con les alumnes en la escuela, y todas ellas fueron aprobadas. Las parejas restantes (6) fueron disueltas, y sus 12 miembros asignades a la condición *de revisión autónoma*.

Etapa de interacción entre pares o revisión autónoma

En la condición de interacción entre pares, ambes niñes de cada pareja debían trabajar juntes. Primero, le participante A ingresaba a la sala, se le daba un cuadernillo, y se le indicaba que debía explicar a su compañere que debían trabajar juntes para hacer un solo dibujo que incluyera todos los términos escritos en el cuadernillo. Mientras tanto, le participante I permanecía afuera con le otre investigadore, con la excusa de responder un cuestionario de preguntas no relacionadas (nombre, edad, mascotas, gustos, etc). Como se comentó en la sección “2.2.1 Introducción”, más allá de esta asimetría de roles, la asimetría de conocimiento no fue comunicada a les participantes.

El cuadernillo incluía los siguientes 14 términos, uno por página: “planeta Tierra”, “luna”, “estrellas”, “sol”, “cielo”, “nubes”, nombre de le participante A, nombre de le participante I, “casa”, “árbol”, “Argentina”, “China (país)”, “persona que vive en China”, “mar”. Le investigadore hacía hincapié al dirigirse a le participante A en que era importante alcanzar un acuerdo antes de dibujar cada elemento del cuadernillo. Todos estos términos habían sido discutidos en la Entrevista 1.

Luego, le participante I ingresaba a la sala y le investigadore, tras recordar a le participante A que debía explicar la tarea a su compañere, dejaba la sala. 45 segundos más tarde regresaba, entregaba el lápiz de dibujo a le participante A, y abandonaba la sala nuevamente. El lápiz no era entregado al comienzo para evitar que les participantes empezaran a dibujar obviando la explicación. Finalmente, cuando les niñes reportaban haber terminado, o cuando hubieran transcurrido 15 minutos (lo que sucediera primero), le investigadore regresaba a la sala y pedía que le explicaran en el orden indicado en el cuadernillo lo que habían dibujado.

Les 12 niñes en el grupo control (C) tenían la misma tarea de hacer un dibujo que incluyera todos los términos en el cuadernillo, pero a solas. En estos casos, los nombres de les participantes A e I en el cuadernillo eran sustituidos por el nombre de le participante control y “un amigo” (o “una amiga”, en caso de participantes de género femenino), respectivamente. En promedio, 10 días pasaron entre la Entrevista 1 y la etapa de interacción entre pares o revisión autónoma (mín = 3d, máx = 17d).

Imagen conflicto 1

En caso de que le o les participantes hubieran dibujado la Tierra redonda, se les mostraba una fotografía de un árbol en un terreno plano y se les daba hasta 9 minutos para pensar y acordar (en ausencia de le investigadore) una respuesta a la cuestión de por qué en la fotografía la Tierra parece plana mientras que en su dibujo es redonda. Al finalizar, le investigadore regresa, (1) repite la pregunta y escucha las respuestas, y finalmente (2) pregunta si, entonces, la Tierra es plana o redonda; siempre asegurando que ambes participantes tengan la posibilidad de expresarse.

Aunque finalmente no analizamos las interacciones y respuestas en torno a esta parte de la tarea (ver apartado *Imagen conflicto 2*, más adelante), incluimos su descripción para garantizar la reproducibilidad de los resultados.

Entrevista 2

Todes les niñes finalmente participaron de una tercera sesión en la que les preguntamos las mismas preguntas que en la Entrevista 1, siguiendo exactamente el mismo procedimiento. Cabe destacar que le entrevistadore, si bien no era ajene a los detalles del estudio, (1) era distinte al que habían tenido en la Entrevista 1, (2) seguía un cuestionario estrictamente guionado, y (3) desconocía si les participantes de la condición de interacción pertenecían al grupo A o I. En promedio, 7 días pasaron entre la etapa de interacción entre pares o revisión autónoma y la Entrevista 2 (mín = 3d, máx = 11d), y 17 entre la Entrevista 1 y la 2 (mín = 8d, máx = 24d).

Esquema de codificación 2 detallado

Uno de los objetivos principales de este trabajo fue investigar cómo las representaciones mentales de la Tierra cambian después de la interacción entre pares, es decir entre las Entrevistas 1 y 2. Sin embargo, las comparaciones estadísticas que se pueden realizar con el Esquema de codificación 1 son limitadas, porque las respuestas de les niñes saturan la escala en este esquema general (es decir, muches niñes tienen puntajes máximos) y por lo tanto su distribución no es normal (ver Figura 2.3B).

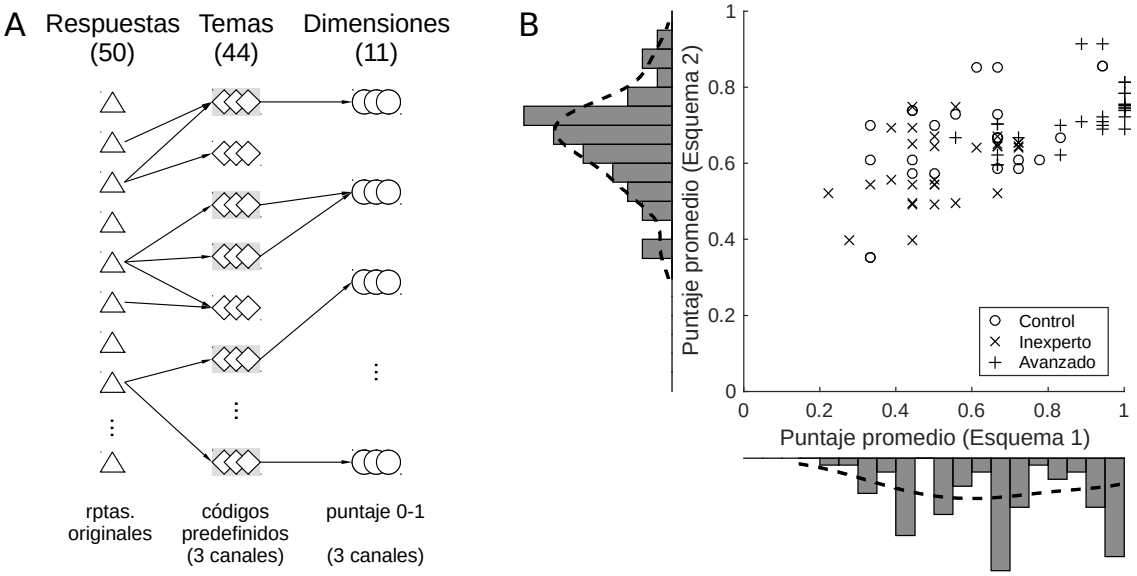


Figura 2.3: *Esquema de codificación 2, detallado*

(A) *Esquema de codificación.* Las respuestas a las preguntas del cuestionario se usan para codificar 44 *temas*, en tres canales cada uno. Algunos de estos *temas* (enmarcados en gris), con códigos predefinidos en dos o más niveles de conocimiento, fueron puntuados de 0 a 1. Finalmente, para evitar que algunos

aspectos del conocimiento estuvieran sobre- o sub-representados (más adelante, en las diferencias de puntaje), los *temas* puntuados que refieren a aspectos similares fueron combinados en 11 *dimensiones* y sus puntajes promediados.

(B) *Comparación con Esquema de codificación 1.* En esta figura se comparan el Esquema de codificación 1 usado para armar las parejas, y el Esquema de codificación 2 empleado para el análisis de datos. Cada punto representa una entrevista (de une participante del grupo *Control*, *Inexperto*, o *Avanzado*), con su puntaje de dimensión promedio del Esquema 2, y su puntaje de eje promedio del Esquema 1. Se encontró una estrecha correlación entre los puntajes de ambos esquemas. El histograma al lado de cada eje representa la frecuencia de los distintos puntajes promedio, con estimación *kernel* de la función de densidad de probabilidad en líneas punteadas.

Para resolver este inconveniente, las respuestas al cuestionario fueron recodificadas utilizando una escala más detallada, de mayor resolución y objetividad. Se identificaron 44 *temas* que podían ser evaluados cuantitativamente a partir de las respuestas de les niñes. Estos temas se refieren a diferentes formas de conocimiento acerca de la Tierra, desde geográficas hasta astronómicas, tales como hacia dónde hay que mirar para ver a la Tierra, cuál es la ubicación de China con respecto a la de Argentina (países ubicados aproximadamente en las antípodas, uno del otro), la posición de la luna con respecto a la Tierra y al sol, la ubicación de las nubes en relación al “cielo”, etc.

Para cada *tema* fueron definidos una serie de códigos. Los videos de las entrevistas de todes les participantes fueron evaluados y codificados en orden aleatorio por una asistenta de investigación ciega al estudio. Para la codificación de una entrevista cualquiera, a cada *tema* le fue asignado uno de los códigos disponibles (para dicho *tema*), según las respuestas provistas por le participante para un conjunto predefinido de preguntas del cuestionario, y siguiendo estrictamente una guía de codificación, con instrucciones detalladas y respuestas ejemplo (ver Anexo 2.2: Guía de codificación). Por ejemplo, las respuestas a las preguntas “Dibujá el planeta Tierra acá” (P09), “¿Así se vería [la Tierra] desde una nave espacial?” (P10), y “¿Cómo se vería desde una nave espacial?” (P11), se usaron para codificar el *tema* “forma abstracta del planeta” (ver Anexo 2.3: Correspondencia entre preguntas y

temas y Anexo 2.4: Temas del Esquema de codificación 2). Cada *tema* podía ser codificado en tres canales, según las respuestas verbales, dibujadas y gestuales, respectivamente (ver Figura 2.3A). Además, para algunos *temas*, sus códigos predefinidos podían ordenarse en *niveles*, de respuestas menos a más avanzadas. Estos *temas* fueron puntuados de 0 a 1 como medida del nivel de conocimiento (ver Anexo 2.4: Temas del Esquema de codificación 2).

El procedimiento experimental garantizó que para cada entrevista al menos 27 de los 44 *temas* fueran codificados. Estos *temas de codificación obligada* correspondían a aspectos explícitamente referidos en el cuestionario (por ejemplo, la forma geométrica de la Tierra). Los restantes 17 *temas de codificación opcional* referían a cuestiones que podían ser aludidas espontáneamente por les participantes, o no. Por lo tanto, podían ser omitidas en entrevistas en las que les niñes no se hubieran manifestado acerca de ese *tema* en particular. Por ejemplo, uno de los *temas de codificación opcional*, “Contradicción entre planeta redondo y plano”, era codificado solamente si le niñe se refería espontáneamente a la aparente contradicción entre la Tierra esférica y el suelo plano, pero esta cuestión no era explícitamente abordada por ninguna de las preguntas del cuestionario.

Imagen conflicto 2

Al igual que al final de la interacción o revisión autónoma, al final de esta etapa se mostró una fotografía, esta vez de una casa en un terreno plano, y se hicieron las mismas preguntas, aunque en este caso sin tiempo para pensarlas a solas. De esta forma quisimos incluir otra forma de analizar los cambios suscitados por la interacción entre pares. Sin embargo, finalmente nos enfocamos en el análisis de las entrevistas solamente.

2.2.2.4 Puntaje de dimensión y nivel de conocimiento

El Esquema de codificación 2 permite una descripción del conocimiento con alta resolución. Sin embargo, algunos aspectos del conocimiento, como la forma de la Tierra, están representados en varios *temas*, mientras que otros en sólo algunos, como la ubicación del mar. Para evitar que algunos aspectos del conocimiento estuvieran sobre- y otros sub-representados en el puntaje general, los *temas* puntuados que refieren a aspectos similares fueron combinados y sus puntajes promediados, dando lugar a 11 *dimensiones* (cada una incluyendo al menos un *tema*

de codificación obligada) y sus correspondientes *puntajes de dimensión*. Estas dimensiones son: ubicación de la Tierra, forma de la Tierra, forma de la luna, ubicación de otros cuerpos celestes, ubicación del cielo, ubicación de le participante en la Tierra, representación y ubicación del país, ubicación de la antípoda, ubicación del mar, exploración de la superficie terrestre, y bordes del planeta (ver Figura 2.3A y "Anexo 2.4: Temas del Esquema de codificación 2" para una descripción más detallada).

La Figura 2.3B muestra que los *puntajes de dimensión* promedio (es decir, el promedio en cada entrevista de los 11 *puntajes de dimensión*) siguen una distribución normal. Además, están estrechamente correlacionados con los *puntajes de eje* promedio (es decir, el promedio en cada entrevista de los 3 *puntajes de eje*) del Esquema de codificación 1, $r(82) = 0{,}65$, $p = 3{,}49E{-}11$. Estos últimos están más íntimamente relacionados con los modelos mentales de la Tierra descriptos previamente (Vosniadou & Brewer, 1992). Un análisis cuantitativo usando la prueba de Shapiro-Wilk y la prueba de *dip* de Hartigan rechazó la normalidad ($W = 0{,}93$, $p = 1{,}27E\text{-}04$) y unimodalidad ($D = 0{,}09$, $p = 6{,}26E\text{-}06$) de los *puntajes de eje* promedio, pero no así las de los *puntajes de dimensión* promedio ($W = 0{,}97$, $p = 0{,}058$; y $D = 0{,}03$, $p = 0{,}840$, respectivamente), que además mostraron menor saturación (Figura 2.3B). Esto indica que los puntajes de dimensión ofrecen una medida del conocimiento conceptual más apropiada para evaluar las diferencias entre niveles de conocimiento en las secciones subsiguientes.

La diferencia de puntaje entre les miembros (participantes A e I) de una pareja (*brecha* de nivel de conocimiento: Δ_{AI}) fue calculada para cada dimensión y luego promediada:

$$(2.1)\quad \Delta_{AI}^{jk} = \left(\sum_{i=1}^{n} d_i^A - d_i^I \right)_{jk} / n \ ,$$

donde Δ_{AI}^{jk} es la diferencia de conocimiento entre les participantes A e I de la *j*-ésima pareja en la entrevista *k*, *n* es el número total de *dimensiones* (n=11), y d_i^A y d_i^I son los puntajes de la i-ésima *dimensión* para les participantes A e I, respectivamente.

Por otra parte, la diferencia de puntaje entre las Entrevistas 1 y 2 de une participante (*cambio* de nivel de conocimiento: Δ_{12}) fue calculada de manera análoga:

$$(2.2)\quad \Delta_{12}^{j} = \left(\sum_{i=1}^{n} d_i^1 - d_i^2 \right)_{jk} / n \ ,$$

donde d_i^1 y d_i^2 son los puntajes de la i-ésima dimensión para las Entrevistas 1 y 2, respectivamente, de le j-ésime participante (ver Figura 2.4). En ambos casos, cada puntaje de dimensión es el promedio del puntaje en los tres canales.

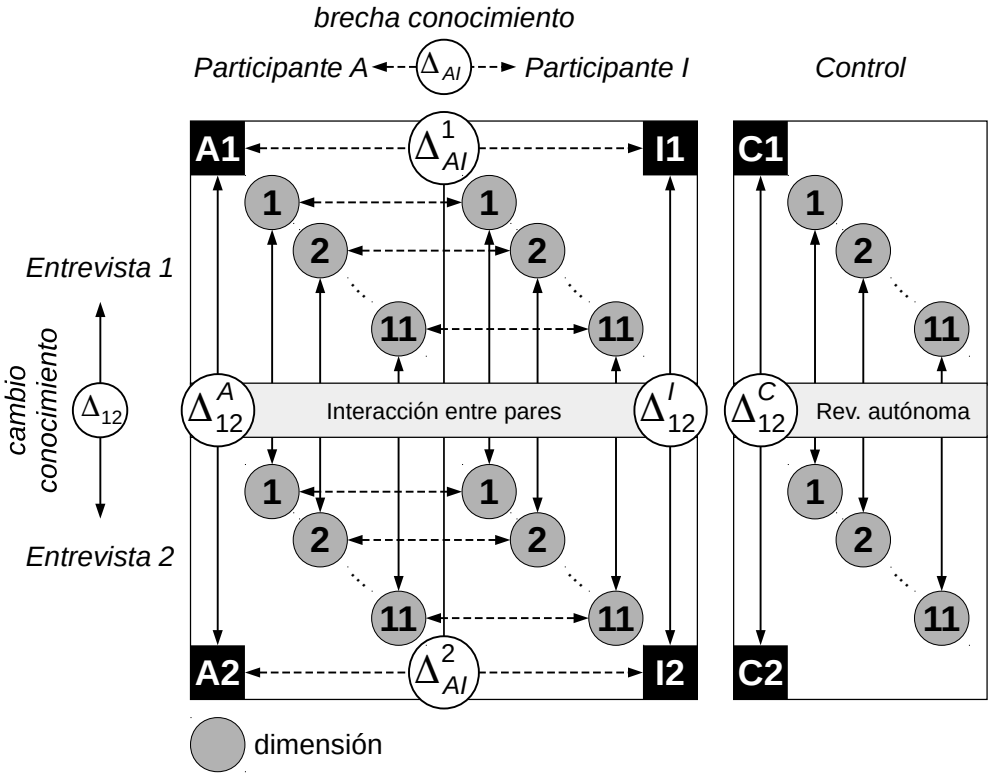


Figura 2.4: *Puntaje de dimensión y nivel de conocimiento*

El nivel de conocimiento se calificó en 11 *dimensiones* por entrevista. Las diferencias entre las entrevistas de une misme participante (*cambio*, Δ_{12}) y entre les miembros de una pareja (*brecha*, Δ_{AI}) fueron calculadas por *dimensión*, y luego promediadas. Δ_{12} : *cambio* de nivel de conocimiento entre las Entrevistas 1 y 2 de une participante, para les participantes A (Δ_{12}^{A}), I (Δ_{12}^{I}) y control (Δ_{12}^{C}); Δ_{AI} : *brecha* de nivel de conocimiento entre les miembros A e I de una pareja, para las Entrevistas 1 (Δ_{AI}^{1}) y 2 (Δ_{AI}^{2}). A1/A2, I1/I2, C1/C2: Entrevista 1 o 2 de participante *Avanzado* (A), *Inexperto* (I), y *Control* (C), respectivamente.

2.2.2.5 Coincidencia por temas y similitud de conocimiento

Para evaluar la similitud de conocimiento entre dos entrevistas –de participantes diferentes o de une participante con sí misme– se comparó la codificación de los *temas*. Concretamente, se consideró que existe coincidencia en un *tema* si fue codificado con exactamente el mismo código en ambas entrevistas para al menos uno de los tres canales: verbal, dibujado o gestual. Códigos diferentes dentro de un mismo nivel de conocimiento no fueron considerados como coincidencia (Figura 2.5).

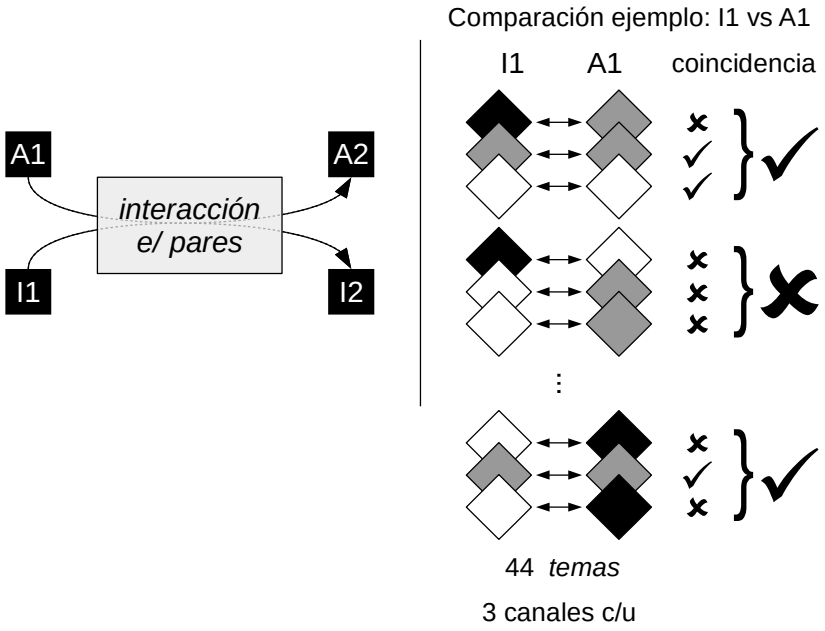


Figura 2.5: *Similitud de conocimiento*

Izquierda. Las respuestas de les participantes de los grupos A e I fueron codificadas en las Entrevistas 1 (A1 e I1) y 2 (A2 e I2).

Derecha. Para evaluar la similitud de conocimiento entre entrevistas –de participantes diferentes o de une misme participante–, se comparó la codificación en cada *tema* y canal. En el ejemplo, se comparan las Entrevistas 1 de les miembros A e I de una pareja (I1A1). Se considera que hay coincidencia en un *tema* si éste fue codificado con exactamente el mismo código en ambas entrevistas en al menos un canal (cada uno de los tres rombos por *tema* en la figura).

2.2.2.6 Cuestionario de comportamiento

Las docentes respondieron un cuestionario de comportamiento para cada une de sus estudiantes participantes. Este cuestionario es una adaptación local basada en las versiones en español (Salinas et al., 2000) y para docentes (Allan et al., 2013) de la forma “muy corta” (Putnam & Rothbart, 2006) del cuestionario de conducta infantil CBQ de Rothbart y cols. (2001), una medida del temperamento establecida para niñes de entre 3 y 8 años. La versión “muy corta” fue elaborada a partir de los tres factores generales que surgen del análisis factorial del cuestionario original: extraversión, afecto negativo y control intencional; que a su vez recuerdan a tres de las cinco dimensiones de personalidad del “Modelo de los cinco grandes” (Digman, 1990): extraversión, neuroticismo, y conciencia, respectivamente.

Por tratarse de una escuela bilingüe, en la que algunes de sus estudiantes son hablantes natives de francés, el cuestionario también indagaba acerca de la habilidad en español y en francés. Además, preguntaba si le estudiante asistía a la actividad extracurricular de astronomía ofrecida por la institución (dos participantes asistían a esta actividad, ningune de les cuales fue excluide de la muestra).

El objetivo de este cuestionario fue revelar aspectos de la personalidad de les participantes que pudieran explicar parte de la variabilidad de los resultados obtenidos. Sin embargo, finalmente concentramos nuestros recursos en el análisis de las entrevistas y aspectos demográficos más generales, como el género y la edad de los participantes.

2.2.2.7 Análisis de datos

El análisis de datos fue realizado en Matlab R2015a para Linux. Se reportan pruebas de dos colas. La prueba de Shapiro-Wilk de normalidad y la prueba *dip* de Hartigan de unimodalidad fueron realizadas en R 3.0.2.

Asimismo, verificamos el cumplimiento de los supuestos para las pruebas estadísticas paramétricas utilizadas.

2.2.3 Resultados

En resumen, el conocimiento de les participantes fue evaluado primero antes de la interacción entre pares, para conocer sus modelos mentales iniciales. Esta

información fue utilizada para agrupar niñes con miradas contrastantes en parejas, tal que une de les miembros poseía una comprensión más acabada (Participante A: *Avanzado*) que le otre (Participante I: *Inexperto*). A continuación, ambes miembros de cada pareja trabajaron juntes en una tarea de argumentación colaborativa. Finalmente, los modelos mentales de todes les niñes fueron evaluados nuevamente después de la etapa de interacción.

2.2.3.1 Cambios de nivel de conocimiento

El conocimiento conceptual de les niñes acerca de la Tierra fue cuantificado codificando sus respuestas verbales, dibujadas y gestuales en 44 *temas* que describen distintos aspectos de este conocimiento (la forma de la Tierra, su ubicación respecto de otros objetos astronómicos, la ubicación del mar, de las nubes y de las personas, etc). A su vez, estos 44 *temas* fueron colapsados en 11 *dimensiones*, ofreciendo un resumen de la representación y el conocimiento de la Tierra que poseen les participantes (ver secciones " Esquema de codificación 2 detallado" y "2.2.2.4 Puntaje de dimensión y nivel de conocimiento" para más detalles sobre el esquema de codificación y sobre la reducción a *dimensiones* relevantes). La diferencia en el nivel de conocimiento, ya sea entre dos miembros de una pareja (*brecha* de nivel de conocimiento: Δ_{AI}) o entre dos evaluaciones (antes y después de la etapa de interacción entre pares o de revisión autónoma) de une misme participante (*cambio* de nivel de conocimiento: Δ_{12}), fue calculada para cada *dimensión* y luego promediada (ver Figura 2.4).

Reducción de la brecha

En promedio, la *brecha* de nivel de conocimiento entre les miembros de una pareja (Δ_{AI}) se redujo después de la interacción, como se muestra con una prueba *t* de Student de muestras apareadas, $t(14) = 2{,}75$, $p = 0{,}016$, $d = -0{,}61$, IC 95% [−1,35, −0,12] (Figura 2.6A).

Naturaleza de esta reducción

Esto pudo haber sido causado por (a) un aumento en el nivel de conocimiento de les participantes del grupo I, lo cual puede sugerir que el trabajo colaborativo podría ser beneficioso para que ocurra un cambio conceptual, (b) una disminución en el nivel

de conocimiento de les participantes del grupo A, lo cual puede sugerir que modelos más avanzados podrían ser cuestionados por la interacción con compañeres más inexpertes, o (c) una combinación de ambos. Para distinguir entre estas posibilidades, se analizó el nivel de conocimiento de cada niñe antes y después de la etapa de interacción. Estos análisis mostraron un *cambio* de nivel de conocimiento positivo y significativo para les miembros más inexpertes de las parejas (Δ^{I}_{12}), $t(14) = 3{,}21$, $p = 0{,}006$, $d = 0{,}83$, IC 95% [0,23, 1,41]. Por el contrario, no se observó un *cambio* de nivel de conocimiento significativo para les miembros más avanzades de las parejas (Δ^{A}_{12}), $t(14) = -0{,}45$, $p = 0{,}662$, $d = -0{,}12$, IC 95% [−0,62, 0,39]. Estos resultados indican que la reducción promedio en la brecha de nivel de conocimiento observada fue debida a un aumento en el nivel de conocimiento de les participantes del grupo I, quienes antes de la interacción poseían representaciones mentales de la Tierra menos avanzadas que las de sus compañeres (Figura 2.6B).

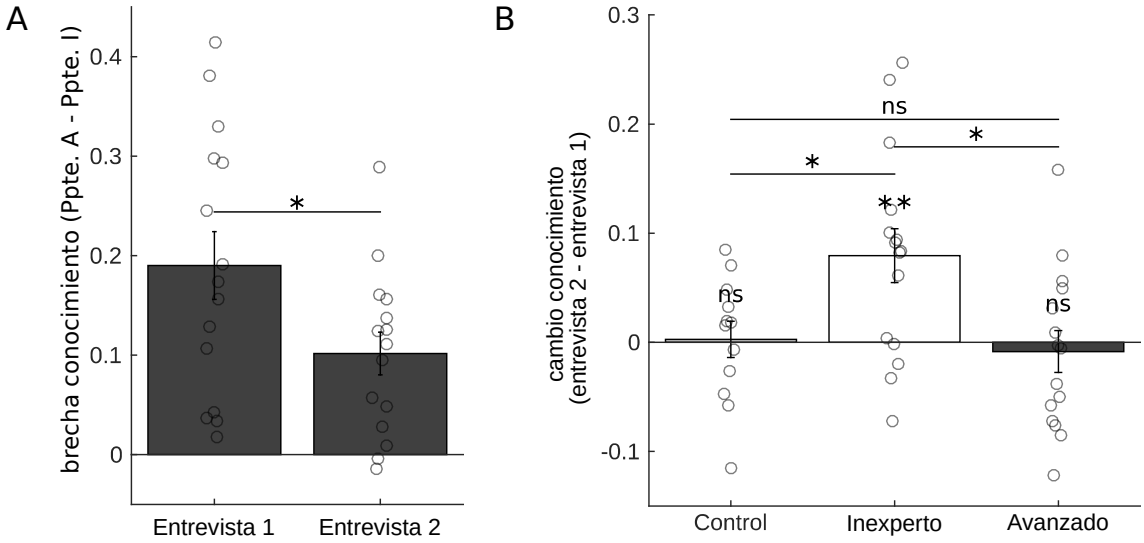


Figura 2.6: *Cambios en los niveles de conocimiento después de la interacción*

(A) En promedio, la *brecha* de nivel de conocimiento entre les miembros de una pareja fue significativamente menor después de la interacción que antes de ésta. Ppte. = *Participante.*

(B) *Cambio* de nivel de conocimiento. Les participantes del grupo I (*Inexperto*) aumentaron significativamente su nivel de conocimiento. En cambio, no se observaron ganancias o pérdidas significativas para les participantes de los grupos A (Aprendiz) o control, ni diferencias significativas entre ambos grupos tampoco.

(*: p < 0,05, **: p < 0.1, ***: p < 0,001, ns: no significativo. Las barras de error

representan el *error estándar de la media*, y los círculos representan datos individuales).

Revisión y repetición como causa posible

Para evaluar si este *cambio* de nivel de conocimiento puede explicarse simplemente por el hecho de haber tenido la oportunidad de considerar el tema nuevamente y/o por ser entrevistade por segunda vez, fue considerado el caso del grupo control. Es decir, el *cambio* de nivel de conocimiento en niñes que revisaron el tema también, pero que lo hicieron a solas sin participar de una tarea colaborativa. Los resultados muestran que no hubo aumento en el nivel de conocimiento en este grupo (Δ^{C}_{12}), $t(11) = 0{,}16$, $p = 0{,}875$, $d = 0{,}04$, IC 95% [−0,53, 0,60] (Figura 2.6B). Esto sugiere fuertemente que el resultado reportado para el grupo I no es atribuible meramente a la repetición del cuestionario o al efecto de reconsiderar el tema durante la tarea de dibujo.

Además, un análisis de la varianza (ANOVA) de una vía mostró un efecto significativo de *grupo* en el *cambio* de nivel de conocimiento entre entrevistas Δ_{12}, $F(2, 39) = 5{,}42$, $p = 0{,}008$, $\eta^2 = 0{,}211$, con pruebas *post hoc* HSD de Tukey mostrando diferencias significativas entre el grupo I (Δ^{I}_{12}: $M = 0{,}08$, $SEM = 0{,}02$) y los grupos control (Δ^{C}_{12}: $M = 0{,}00$, $SEM = 0{,}02$, $p = 0{,}042$) y A (Δ^{A}_{12}: $M = -0{,}01$, $SEM = 0{,}02$, $p = 0{,}011$). Por su parte, los grupos control y A no evidenciaron diferencias significativas entre ellos ($p = 0{,}928$).

Punto de inicio como causa posible

Como las parejas estaban formadas por une miembro con un puntaje menor que le otre, los puntajes de la Entrevista 1 no están igualmente distribuidos entre grupos. En efecto, un ANOVA de una vía muestra diferencias significativas en los puntajes promedio iniciales (Entrevista 1) entre grupos, $F(2, 39) = 10{,}2$, $p = 2{,}69\text{E}{-}04$, $\eta^2 = 0{,}327$. En particular, el promedio de puntajes del grupo I fue el más bajo. Es posible que este puntaje menor haya sido la razón detrás de los *cambios* de nivel de conocimiento observados en este grupo. Sin embargo, si bien las comparaciones con prueba HSD de Tukey muestran diferencias significativas esperables entre los

grupos I (M = 0,53, SEM = 0,04) y A (M = 0,72, SEM = 0,02, p = 1,68E−04), no se encuentran diferencias significativas entre el grupo control (M = 0,61, SEM = 0,03) y los grupos I (p = 0,169) o A (p = 0,053). Además, también hicimos un análisis restringido que sólo tuvo en cuenta participantes cuyo puntaje en la Entrevista 1 se encontraba entre 0,45 y 0,75, un intervalo en el que participantes de ambos grupos control e I se encuentran similarmente distribuidos. Nuevamente en este rango, como fue observado con el conjunto de datos completo, les participantes del grupo I (n = 12) mostraron ganancias de nivel de conocimiento significativas, $t(11)$ = 2,71, p = 0,020, d = 0,88, IC 95% [0,20, 1,54], mientras que les participantes del grupo control (n = 11) no, $t(10)$ = 0,08, p = 0,935, d = 0,02, IC 95% [−0,57, 0,61].

En resumen estos resultados muestran que la representación conceptual colectiva de las parejas acerca de la Tierra progresó. Esto fue consecuencia de que las representaciones mentales menos avanzadas se aproximaron a modelos más sofisticados, sin evidencia de que haya habido algún perjuicio para les miembros con nivel de conocimiento más avanzado.

2.2.3.2 Cambios de similitud de conocimiento

Los resultados descriptos arriba muestran que el nivel de conocimiento aumentó después de interactuar con une par más avanzade. Esto tiene dos explicaciones posibles. Una posibilidad fundamentalmente constructivista es que el proceso de reflexión con une par más avanzade haya llevado a les niñes a cuestionarse –en un sentido amplio y general– aspectos de sus propias representaciones conceptuales. Otra posibilidad es que les niñes más inexpertes hayan adoptado hechos y elementos específicos de las representaciones mentales de sus compañeres de interacción (Smith et al., 2009). Si esta última posibilidad fuese el caso, les participantes del grupo I no sólo habrían aumentado sus niveles de conocimiento, sino que además expresarían representaciones similares a las de sus compañeres más avanzades. Para explorar este aspecto, cuantificamos el grado de similitud de conocimiento entre ambes miembros de cada pareja, en un espacio de conocimiento multidimensional, antes y después de la interacción.

La similitud de conocimiento se puede expresar como la inversa de la distancia de Hamming (entendida como el número de dimensiones en las que difieren dos

vectores) entre dos vectores de representación conceptual. Así expresado, este índice mide en cuántos de los 44 *temas* ambas representaciones vectoriales son idénticas. Como se explica en detalle en la sección “2.2.2.5 Coincidencia por temas y similitud de conocimiento”, dos *temas* coinciden en una comparación vectorial (ya sea entre dos miembros de una pareja, o entre dos entrevistas de une misme participante) si fueron asignades exactamente el mismo código predefinido en al menos uno de los tres canales posibles: verbal, dibujado o gestual (Figura 2.5). Por ejemplo, si el *tema* “T01: Hacia dónde mirar para ver el planeta Tierra” fue codificado con “T01.00: Desde aquí hacia arriba o hacia el cielo” en el canal gestual para ambos casos (ambes miembros de una pareja o ambas entrevistas de une misme participante), diremos que el *tema* coincide entre ambos casos. Pero si fue codificado con “T01.00: Desde aquí hacia arriba o hacia el cielo” en un caso, y con “T01.10: Hacia o en el espacio” en el otro, diremos que no hay coincidencia.

Aumento de la similitud

Primero, comparamos el número de coincidencias por *tema* entre les miembros de una pareja antes (I1A1) y después (I2A2) de la interacción entre pares (ver Figura 2.5). Una prueba *t* de Student de muestras apareadas muestra que el número promedio de coincidencias por *tema* entre pares aumentó significativamente después de la tarea colaborativa, $t(14) = 3{,}41$, $p = 0{,}004$, $d = 0{,}99$, IC 95% [0,23, 1,75] (Figura 2.7A).

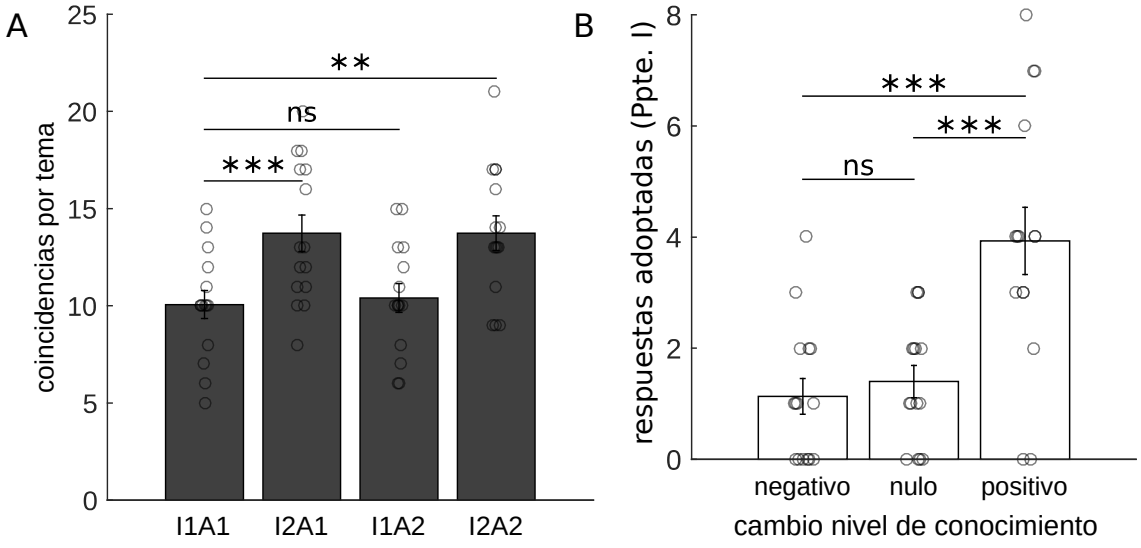


Figura 2.7: *Cambios en la similitud de conocimiento*

(A) *Coincidencias entre les miembros de una pareja.* Les participantes del grupo I y

sus respectives compañeres coincidieron en significativamente más *temas* después de interactuar (I2A2) que antes de hacerlo (I1A1). Particularmente, la *adopción de respuestas* (es decir, la modificación de las respuestas de une participante para asemejarse a las de su compañere) fue significativa para les participantes del grupo I ($\Delta_{I2A1\text{-}I1A1}$), pero no para les del grupo A ($\Delta_{I1A2\text{-}I1A1}$). IxAy: número de coincidencias entre le miembro I de una pareja en la Entrevista x, y le miembro A de la misma pareja en la Entrevista y, con x/y: 1 o 2.

(B) *Cambios* de nivel de conocimiento motivados por *adopción de respuestas*. Una respuesta adoptada se define como un *tema* que coincide entre la Entrevista 2 de une participante y la Entrevista 1 de su par, pero que no coincide entre sus Entrevistas 1 antes de la interacción. En esta figura, las respuestas adoptadas se dividen en grupos en función del carácter (positivo, negativo o nulo) del cambio de nivel de conocimiento producido. En les participantes del grupo I (Ppte. I), las respuestas adoptadas condujeron significativamente a más ganancias en el nivel de conocimiento que a pérdidas o cambios nulos.

(*p < 0,05, **p < 0.1, ***p < 0,001, ns: no significativo. Las barras de error representan el *error estándar de la media*, y los círculos representan datos individuales).

Naturaleza de este aumento

Este mayor grado de similitud puede ser el resultado de (i) les participantes más inexpertes (I) asimilándose a las representaciones mentales de sus pares más avanzades (A); (ii) el proceso inverso, con les participantes más avanzades asimilándose a les más inexpertes; o (iii) ambos procesos, con las representaciones de les participantes convergiendo en algún punto intermedio. Para distinguir entre estas posibilidades, se cuantificó el número de respuestas que les participantes adoptaron de sus pares. Nos referiremos a esta magnitud como adopción de respuestas: $a^I = \Delta_{I2A1-I1A1}$, para el grupo I, y $a^A = \Delta_{I1A2-I1A1}$, para el grupo A; donde $I2A1$ e $I1A2$ son el número de coincidencias por *tema* entre la Entrevista 2 de le participante del grupo I o A, respectivamente, y la Entrevista 1 de su par (ver Figura 2.5). Pruebas *t* de Student con muestras apareadas muestran que la adopción de respuestas es significativa para el grupo I, $t(14) = 4{,}91$, $p = 2{,}32E{-}04$, $d = 1{,}51$, IC 95% [0,70, 2,32], pero no para el grupo A, $t(14) = 0{,}42$, $p = 0{,}680$, $d = 0{,}11$, IC 95% [−0,61, 0,83], (Figura 2.7A).

Número de temas y canales codificados

Dado que, por la naturaleza de las respuestas de les participantes, no todos los *temas* o canales fueron codificados en todas las entrevistas, realizamos una serie de análisis adicionales (ver Tabla 2.1) para asegurar que estas diferencias en el número de *temas* y canales codificados no hubieran tenido un efecto sobre las diferencias reportadas en el número de coincidencias por *temas*.

Análisis adicional	Condición		
	Ponderada	Sin opcionales	Ambas entrevistas
0 *	No	No	No
1	Sí	No	No
2	No	Sí	No
3	No	No	Sí

Tabla 2.1: *Condiciones de análisis adicionales*

Ponderada: el número de coincidencias por tema se pondera según el número total de *temas* codificados (para al menos une de les miembros de la pareja). *Sin opcionales:* se excluyen los *temas de codificación opcional*; es decir, aquellos que no corresponden a aspectos explícitamente referidos en el cuestionario, sino que podían ser espontáneamente aludidos o no por les participantes.[4] *Ambas entrevistas:* para cada tema y sujeto, se consideran sólo los canales empleados en ambas entrevistas.

(*) Análisis 0: el análisis de base, como se reporta en los apartados anteriores de esta sección.

Estos análisis complementarios muestran (ver Tabla 2.2) que las respuestas adoptadas por les participantes del grupo I no corresponden (o no corresponden exclusivamente) a *temas* o *canales en* que estes participantes no se hubieran expresado antes de la interacción, sino que efectivamente implican la *modificación* de sus respuestas luego de la tarea colaborativa.

4 Por error, los *temas de codificación opcional* T36 y T37 no fueron excluidos en la condición de análisis adicional “sin opcionales”.

Análisis adicional	Coincidencia por temas intra pareja (entre sujetos)													
	I1A1		I2A1 (1)				I1A2 (1)				I2A2 (1)			
	M	*SEM*	*M*	*SEM*	*t*(14)	*p*	*M*	*SEM*	*t*(14)	*p*	*M*	*SEM*	*t*(14)	*p*
0 *	10.07	0.72	13.73	0.94	4.91	< .001	10.40	0.74	0.42	.680	13.73	0.89	3.41	.004
1	0.28	0.02	0.37	0.03	4.39	.001	0.28	0.02	0.11	.916	0.37	0.03	3.03	.009
2	9.80	0.71	13.07	0.86	4.62	< .001	10.20	0.67	0.52	.613	13.20	0.85	3.22	.006
3	9.87	0.71	12.80	0.95	4.11	.001	9.93	0.64	0.08	.934	12.67	0.75	3.33	.005

Tabla 2.2: *Resultados de análisis adicionales*

IxAy (ej., I1A1): número de coincidencias por tema entre la Entrevista *x* del participante del grupo I, y la Entrevista *y* del participante del grupo A, con x/y = 1 o 2.

(*) Análisis 0: el análisis de base, como se reporta en el texto principal.

(1) I2A1, I1A2, e I2A2 se comparan frente a I1A1.

En resumen, por un lado los resultados de la sección anterior muestran que les participantes más inexpertes aumentaron sus niveles de conocimiento, evaluado en términos de su cercanía con el conocimiento científico. Los resultados de esta sección muestran que la similitud de conocimiento entre les participantes aumentó después de la interacción, producto de la adopción por parte de les participantes más inexpertes de las respuestas de sus pares más avanzades.

2.2.3.3 Adopción de respuestas y cambios de nivel de conocimiento

Los resultados hasta aquí muestran que, después de una interacción colaborativa entre pares, (i) les participantes menos avanzades aumentaron sus niveles de conocimiento (Figura 2.6B), y que (ii) adoptaron respuestas de sus pares más avanzades, asimilándose a sus representaciones mentales (Figura 2.7A). Sin embargo, la adopción de respuestas puede haber llevado a cambios de nivel de conocimiento positivos, negativos, o nulos (sin cambios). Para verificar si efectivamente los niveles de conocimiento aumentaron producto de la adopción de respuestas correctas, se analizó la naturaleza de estas adopciones.

Como vimos anteriormente, una respuesta adoptada de une par es un *tema* que coincide entre la Entrevista 2 de le participante bajo consideración, y la Entrevista 1 de su par (es decir, entre I2 y A1, o entre A2 e I1), pero que no coincidía antes de

interactuar, en la Entrevista 1 (es decir, entre I1 y A1; ver Figura 2.5). Sólo para el grupo I, la adopción de respuestas condujo a cambios de nivel de conocimiento positivos y significativos. Un ANOVA de una vía muestra diferencias significativas en el número promedio de respuestas adoptadas en función del tipo de cambio de nivel (positivo, negativo o nulo) producido, $F(2, 42) = 12{,}9$, $p = 4{,}18E{-}05$, $\eta^2 = 0{,}37$ (Figura 2.7B). Comparaciones *post hoc* con prueba *HSD* de Tukey muestran que el número de respuestas adoptadas que produjeron cambios positivos de nivel ($M = 3{,}93$, $SEM = 0{,}61$) es significativamente mayor que el número de adopciones que produjeron cambios negativos ($M = 1{,}13$, $SEM = 0{,}32$, $p = 1{,}09E{-}04$) o nulos ($M = 1{,}40$, $SEM = 0{,}29$, $p = 4{,}31E{-}04$). Por el contrario, no se encontraron diferencias significativas en el número de respuestas adoptadas en función del tipo de cambio de nivel producido para el grupo A, $F(2, 42) = 0{,}67$, $p = 0{,}518$, $\eta^2 = 0{,}03$.

El análisis fue repetido después de ponderar el número de respuestas adoptadas por el total de *temas* codificados, para evitar que participantes con mayor número de *temas* codificados sesgaran los resultados, y obtuvimos los mismos resultados. Concretamente, un ANOVA de una vía muestra diferencias significativas en el número promedio de respuestas adoptadas (relativo al total de temas codificados) en función del tipo de cambio de nivel de conocimiento (positivo, negativo o nulo) producido para el grupo I, $F(2, 42) = 13{,}4$, $p = 3{,}23E{-}05$. Comparaciones post hoc con prueba HSD de Tukey muestran que el número (relativo) de respuestas adoptadas que produjeron cambios positivos de nivel ($M = 0{,}10$, $SEM = 0{,}02$) es significativamente mayor que el número (relativo) de adopciones que produjeron cambios negativos ($M = 0{,}03$, $SEM = 0{,}01$, $p = 8{,}42E{-}05$) o nulos ($M = 0{,}04$, $SEM = 0{,}01$, $p = 3{,}62E{-}04$). Nuevamente, las diferencias no fueron significativas para el grupo A, $F(2, 42) = 0{,}60$, $p = 0{,}555$.

Estos resultados sugieren que les participantes del grupo I adoptaron significativamente más respuestas correctas de sus pares más avanzades del grupo A. Esto habría conducido a los cambios positivos de nivel de conocimiento observados. Por otra parte, las (pocas) respuestas que les participantes más avanzades adoptaron de sus pares más inexpertes no condujeron a cambios de nivel de conocimiento en promedio.

2.2.3.4 Efecto en los modelos mentales de la Tierra

En la sección “2.2.3.1 Cambios de nivel de conocimiento mostramos que la representación conceptual de la Tierra de las parejas progresó, y que esto fue consecuencia de que les niñes con una representación menos avanzada (grupo I) se acercaron a una más avanzada. Este resultado se obtuvo colapsando el conocimiento conceptual en 11 dimensiones para estimar su nivel.

Las preguntas en las entrevistas evalúan aspectos diferentes de la representación conceptual de la Tierra. Para investigar si algunos *temas* son más propensos al cambio que otros, fueron agrupados alternativamente en tres *dominios*: *Cielo, Suelo,* y *Exploración* (ver “Anexo 2.4: Temas del Esquema de codificación 2”). Algunos *temas* involucran nuestro planeta y sus alrededores: su forma y ubicación en el universo, y cómo se relaciona con otras entidades relativas al cielo (es decir, astronómicas y atmosféricas), como la luna, el sol o las nubes. Nos referiremos a estos *temas* como parte del dominio *Cielo.* Otros se relacionan con la disposición de los objetos, como edificios, personas y continentes, en la superficie terrestre, y las relaciones entre ellos. Nos referiremos a este dominio como *Suelo*. Finalmente, un tercer conjunto de *temas* se refieren a cómo el planeta se revelaría ante une exploradore que navega su superficie, qué obstáculos encontraría, si se toparía o no con un borde, etc. Nos referiremos a este dominio como *Exploración*.

El cambio de nivel de conocimiento observado en el grupo I después de la interacción –progresando hacia representaciones más avanzadas– se debe fundamentalmente a cambios en representaciones relativas al dominio *Suelo*. Así surge de pruebas *t* independientes que muestran que el cambio de nivel de conocimiento en el grupo I es significativo solamente para el dominio *Suelo*, $t(14) = 3{,}46$, $p = 0{,}004$, mientras que no lo es en los dominios *Cielo,* $t(14) = 0{,}69$, $p = 0{,}499$, y *Exploración*, $t(14) = 0{,}00$, $p = 1{,}00$. No se observaron cambios significativos en ningún dominio para los grupos control y A (ver Tabla 2.3). Esto además descarta que la ausencia de cambio reportada para estos dos grupos sea consecuencia de compensación entre cambios positivos y negativos para los distintos dominios.

Dominio	Cambio de nivel											
	Control				Grupo I				Grupo A			
	M (1E-2)	*SEM* (1E-2)	*t*(11)	*p*	*M* (1E-2)	*SEM* (1E-2)	*t*(14)	*p*	*M* (1E-2)	*SEM* (1E-2)	*t*(14)	*p*
Todos (1)	0.27	1.65	0.16	.875	7.94	2.47	3.21	.006	-0.86	1.92	-0.45	.662
Cielo	4.17	4.66	0.89	.391	2.75	3.96	.69	.499	-4.94	2.60	-1.90	.078
Suelo	1.90	2.88	0.66	.523	20.90	6.05	3.46	.004	-2.30	2.02	-1.14	.275
Exploración	-4.17	5.08	-0.82	.429	0.00	5.63	0.00	1.00	-1.11	8.37	-0.13	.896

Tabla 2.3: *Cambio de nivel de conocimiento para cada grupo de temas o dominio*

(1) Todos: valores reportados en la sección “2.2.3.1 Cambios de nivel de conocimiento”, con los *temas* agrupados en 11 dimensiones en vez de en 3 dominios.

No se encontraron diferencias significativas para ninguno de los dominios en el puntaje de Entrevista 1 entre ambos grupos de interacción entre pares (grupos I y A) y el grupo control (ver Tabla 2.4), sugiriendo nuevamente que las diferencias observadas entre grupos no se deben simplemente a distintos puntos de inicio antes de la interacción entre pares o revisión autónoma.

Dominio	Puntajes iniciales (Entrevista 1)										
	Control		Grupo I		Grupo A		ANOVA		Prueba HSD de Tukey, *p*-valor		
	M	*SEM*	*M*	*SEM*	*M*	*SEM*	F(2, 39)	*p*	Control vs Grupo I	Control vs Grupo A	Grupo I vs Grupo A
Todos (1)	0.61	0.03	0.53	0.04	0.72	0.04	10.23	< .001	.169	.053	< .001
Cielo	0.73	0.05	0.69	0.03	0.82	0.03	3.53	.039	.730	.220	.034
Suelo	0.63	0.06	0.48	0.07	0.81	0.02	10.36	< .001	.124	.070	.001
Exploración	0.38	0.03	0.27	0.05	0.41	0.06	2.28	.116	-	-	-

Tabla 2.4: *Puntaje inicial en Entrevista 1 por grupo de temas o dominio*

(1) Todos: valores reportados en la sección “2.2.3.1 Cambios de nivel de conocimiento”, con los *temas* agrupados en 11 dimensiones en vez de en 3 dominios.

Para ayudarnos a entender qué podría significar este cambio de nivel de conocimiento en el dominio *Suelo* en término de los modelos mentales de la Tierra descriptos previamente en la literatura (Mali & Howe, 1979; J. Nussbaum, 1979; Sneider & Pulos, 1983; Vosniadou & Brewer, 1992), empleamos un análisis de regresión múltiple para evaluar si los puntajes de eje del Esquema de codificación 1, más íntimamente relacionados con estos modelos previamente descriptos, predicen significativamente el puntaje promedio en el dominio *Suelo*. Usando un ajuste por

cuadrados mínimos, la correlación múltiple identificada es significativa, con los puntajes de eje explicando un 53% del puntaje promedio en el dominio *Suelo*, $R^2 = 0{,}53$, $F(4, 80) = 30{,}1$, $p = 4{,}11E{-}13$. En particular, el puntaje en el eje dual/no dual es un predictor significativo del puntaje promedio en el dominio *Suelo* ($\beta = 0{,}45$, $p = 1{,}61E{-}12$), mientras que los puntajes en los ejes hueco/no hueco y discoidal/esférico no ($\beta = 0{,}05$, $p = 0{,}365$, y $\beta = 0{,}09$, $p = 0{,}061$, respectivamente).

En conjunto estos resultados sugieren que un cambio conceptual de modelos duales a no duales podría haber ocurrido para les participantes del grupo I después de la interacción con une par más avanzade. En cambio, este proceso no parece haber sido promovido en los participantes del grupo A, o por simple revisión del material en los participantes del grupo control.

2.2.4 Discusión

El objetivo principal de este estudio era comprender la construcción colectiva de conocimiento, y los procesos de desarrollo correspondientes en les individues, en particular cuando les niñes se enseñan entre elles, posiblemente sosteniendo e intercambiando argumentos a favor de modelos conceptuales diferentes. La hipótesis era que el conflicto sociocognitivo entre niñes con miradas contrastantes sobre la Tierra, posiblemente elaborado a través de la argumentación colaborativa, ayudaría a promover la construcción colectiva de conocimiento, acelerando procesos de cambio conceptual.

2.2.4.1 La interacción aumenta los niveles de conocimiento

Primero, la evaluación del conocimiento de les participantes confirmó la amplia diversidad de representaciones mentales descripta en trabajos previos (Blown, 2013; Mali & Howe, 1979; J. Nussbaum, 1979; Saçkes, 2015; Sneider & Pulos, 1983; Vosniadou & Brewer, 1992). Esta diversidad sentó las bases apropiadas para realizar el estudio, ya que permitió establecer conflictos sociocognitivos aprovechando diferencias de conocimiento previas, sin necesidad de introducirlas artificialmente.

Por otra parte, cuando les niñes interactuaron colaborativamente con une par más avanzade, aumentaron significativamente sus niveles de conocimiento al cambiar sus respuestas para asemejarlas a las de sus compañeres. En cambio, estos aumentos de nivel de conocimiento no se observaron en participantes que revisaron el tema a solas, sin interactuar con une par. Este hallazgo sugiere que la construcción de conocimiento fue efectivamente promovida por la interacción entre pares, y no simplemente por la revisión autónoma de los conceptos o por la repetición de la entrevista.

Comprensión vs imitación

Cabe preguntarse si estos resultados manifiestan una comprensión conceptual verdadera, o si por el contrario se explican por simple repetición e imitación. A este respecto, por un lado es relevante señalar que la Entrevista 2 o etapa de reevaluación fue análoga, pero no idéntica, a la etapa de interacción entre pares o revisión autónoma. No sólo el marco era diferente (un cuestionario vs una tarea de dibujo), sino que además el contexto social lo era también (guiado por une adulte vs colaborativo/solitario). Este isomorfismo, de la etapa de interacción entre pares o revisión autónoma a la etapa de reevaluación, supone la aplicación de los mismos principios o conceptos para resolver tareas diferentes (Smith et al., 2009). Esta *transferencia* hace que sea menos probable que les niñes estuvieran simplemente imitando a sus pares sin comprender realmente (Smith et al., 2009). Además, por otra parte, los cambios de conocimiento no fueron evaluados inmediatamente después de la etapa de interacción entre pares o revisión autónoma, sino con una diferencia de una semana en promedio. El hecho de que los cambios de conocimiento fueran preservados tras este período es compatible con la comprensión de los conceptos en oposición a la simple repetición (Asterhan & Schwarz, 2007). Mayores diferencias entre las tareas, y períodos más largos entre la etapa de interacción entre pares y la de reevaluación ayudarían a distinguir aún más entre estas posibilidades.

2.2.4.2 Efecto tutor

De acuerdo con resultados previos en el campo de la enseñanza entre pares, nuestras hipótesis incluían que les niñes más avanzades se beneficiarían también de la interacción colaborativa (P. A. Cohen et al., 1982; Duran, 2016; Roscoe & Chi,

2008; K. Topping, 2015, 2017), en especial teniendo en cuenta que, por la reducida brecha con sus pares menos avanzades, era de esperar que se enfrentarían a desafíos cognitivos también (K. Topping, 2015). Sin embargo, no encontramos evidencia a favor de esta hipótesis. Una posible explicación es que, mientras el puntaje inicial (en la Entrevista 1) de les niñes menos avanzades del grupo I era variable, de bajo a medio, el grupo A de niñes más avanzades tenía puntajes iniciales más altos. Esto podría haber creado un *efecto techo*, limitando la observación y cuantificación de cambios en el nivel de conocimiento. Además, mientras el cambio promedio de nivel en el grupo A fue nulo, se observó una posible tendencia hacia ganancias mayores en aquelles que habían interactuado con les participantes más avanzades del grupo I (Figura 2.8). Esta tendencia puede sugerir que les niñes más avanzades se podrían beneficiar con la argumentación colaborativa si se enfrentan a desafíos cognitivos mayores a través de interacciones que involucren brechas de conocimiento menores entre les miembros de la pareja (K. Topping, 2015). Sin embargo, esta tendencia no es significativa y, por lo tanto, no podemos concluir que les niñes del grupo A podrían beneficiarse de la interacción entre pares.

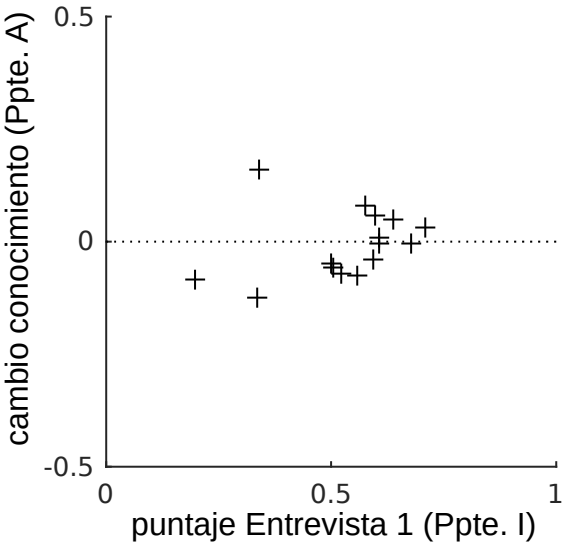


Figura 2.8: *Cambio de nivel de conocimiento como función de la brecha*

Cambio de nivel de conocimiento en participantes del grupo A (Ppte. A) en función del puntaje en la Entrevista 1 de su compañere del grupo I (Ppte. I). Mientras el cambio promedio de nivel de conocimiento en el grupo A fue nulo, descriptivamente se observa una posible tendencia hacia ganancias mayores en aquelles que habían

interactuado con les participantes más avanzades del grupo I, $r(13) = 0,27$, $p = 0,331$.

No obstante, cabe destacar que tampoco se observó una disminución significativa del nivel de conocimiento producto de esta interacción con une par menos avanzade. En este estudio, el desacuerdo social y la argumentación en niñes condujo a la construcción de conocimiento más avanzado: le participante más avanzade ayudó a promover el cambio conceptual en le menos avanzade, sin efectos negativos sobre su propio conocimiento. Desde la perspectiva de la corriente del *conocimiento como teoría* (Vosniadou et al., 2009), este hallazgo es compatible con la idea Piagetiana de que los conceptos se vuelven *verdades necesarias* una vez que han sido comprendidos plenamente (S. A. Miller & Brownell, 1975), y ayuda a disipar la noción de que la interacción entre pares podría ser contraproducente para les niñes que deben colaborar con pares menos avanzades (Kalkowski, 1995).

2.2.4.3 Interacción colaborativa y consenso

Los resultados de este estudio muestran que la brecha de conocimiento dentro de las parejas se redujo porque el nivel de conocimiento de les miembros menos avanzades aumentó, y no porque el de les más avanzades hubiera disminuido. Esto sugiere que les niñes no convergieron en un acuerdo promedio entre sus puntos de vista originales, sino que el consenso se ubicó en algún punto inclinado hacia la posición de le miembro más avanzade. Esta inclinación pudo estar dada por distintas heurísticas, tales como (a) la confianza percibida de le otre con respecto a su punto de vista, (b) indicios de autoridad, posiblemente provistos por la asignación asimétrica de roles (Cialdini, 2006; Harvey & Hays, 1972), o (c) argumentos ofrecidos por ambas partes.

El hecho de que los aumentos de nivel de conocimiento fueran preservados tras un período de una semana en promedio, es compatible con el tipo de aprendizajes que derivan de la elaboración del conflicto sociocognitivo a través de la argumentación colaborativa (Asterhan & Schwarz, 2007). El compromiso de les participantes con la tarea de interacción, el hecho que debían dar una sola respuesta conjunta (E. M. Nussbaum, 2008), y la asimetría en el nivel de conocimiento, pueden ser algunas de las razones de por qué la colaboración libre podría haber derivado espontáneamente

–sin instrucción explícita– en argumentación colaborativa. Un estudio detallado de las interacciones entre pares en sí mismas ayudaría a confirmar si el conflicto sociocognitivo fue efectivamente elaborado a través de la argumentación colaborativa, y a establecer las condiciones bajo las cuales la colaboración libre promueve el cambio conceptual.

Cabe recordar que en este estudio se introdujo también una sutil asimetría de roles al dar sistemáticamente a le miembro más avanzade las instrucciones de lo que debían hacer en pareja. Aunque esta asimetría fue pequeña y no implicaba explícitamente que une participante supiera más que le otre, o que une debía enseñar a le otre, su efecto sobre la naturaleza de los consensos alcanzados por las parejas no puede descartarse.

2.2.4.4 Del pensamiento aristotélico al copernicano

Los *temas* detrás de los aumentos de nivel de conocimiento reportados habrían sido principalmente aquellos que refieren a aspectos relacionados con la ubicación de los objetos en la superficie terrestre. Éstos, a su vez, se asocian con el eje dual/no-dual del Esquema de codificación general 1 (ver apartado *Esquema de codificación 1 general* de la sección “2.2.2.3 Diseño experimental”). Este resultado sugiere que la interacción entre pares podría haber favorecido la revisión, por parte de les participantes menos avanzades, de sus modelos de Tierra dual, que como vimos anteriormente son los modelos sintéticos alternativos más sencillos y habitualmente los primeros en aparecer a lo largo del camino de desarrollo conceptual.

Por lo tanto, los cambios de conocimiento observados no sólo indican que les niñes menos avanzades se beneficiaron de la interacción entre pares aumentando su nivel de conocimiento. También sugieren que la interacción podría haber promovido este aumento ayudándoles a reinterpretar que los objetos sin sostén caen, permitiéndose conceptualizar a la Tierra como un objeto suspendido en el espacio, y cambiando así de un modelo dual a uno no-dual. Esto es compatible con procesos de desarrollo conceptual profundos, en contraste con el simple aprendizaje de hechos puntuales, y constituye un paso importante hacia reconceptualizar la Tierra como un objeto astronómico (Vosniadou, 1994). Esta recategorización ontológica, que precede a la comprensión plena de la Tierra como un planeta esférico (Vosniadou et al., 2009), habitualmente ocurriría entre el 3ro y 5to grado de escuela primaria (Vosniadou &

Skopeliti, 2014). Así, en otras palabras, le par más avanzade habría guiado a su compañere a entender aquello que ya está listo para comprender, combinando las visiones en apariencia opuestas de un objeto suspendido en el espacio, y la de la superficie aparentemente plana en la que vivimos. De esta manera, la interacción entre pares podría catalizar el cambio conceptual, que habría ocurrido de todos modos, pero más lentamente.

En línea con propuestas previas para reconciliar las miradas cognitivista y socioculturalista del cambio conceptual, el paradigma empleado en este estudio satisface tres requerimientos propuestos para que ocurra este salto conceptual (Heyd-Metzuyanim & Schwarz, 2017). Por un lado, (1) las asimetrías de nivel de conocimiento habrían provisto discursos alternativos a les niñes menos avanzades. Además, (2) el hecho de que estas asimetrías no fueran explícitas podría haberse traducido en niveles más altos de igualdad, que a su vez podrían haber ayudado a poner el conflicto en evidencia, en vez de alentar a que les niñes más avanzades impusieran su conocimiento sobre sus compañeres. Finalmente, (3) este aspecto implícito de las asimetrías de rol y de conocimiento podría haber acercado las posiciones de les miembros de las parejas, y contribuido a mantener niveles de mutualidad altos, típicos de las interacciones de enseñanza entre pares (Damon & Phelps, 1989), lo que probablemente ayudó a les menos avanzades a reconocerse capaces de dar el salto conceptual necesario.

2.2.4.5 Impacto pedagógico

Los resultados de este trabajo proveen evidencia empírica y sugieren formas específicas en las que la interacción entre pares se puede usar efectivamente para promover cambios conceptuales en contextos escolares, en un dominio del conocimiento tan representativo de las crisis de posverdad y de rechazo del conocimiento científico que caracterizan nuestra época. Sobre los hombros de la identificación previa de los conceptos que tienen les niñes sobre la Tierra, este estudio se ha enfocado en los procesos de aprendizaje en este dominio, y en cómo promoverlos (Hobson et al., 2010). Esto podría eventualmente permitir una interpretación y transferencia al campo educativo, pedagógico y de diseño de políticas públicas.

La argumentación es clave para el aprendizaje de las ciencias. Sin embargo, la enseñanza en ciencia muchas veces se destaca por la falta de argumentos (Osborne, 2010). Además, se necesitan prácticas de enseñanza específicas para favorecer los procesos de cambio conceptual involucrados (Hadzigeorgiou, 2015), que a veces las explicaciones de les instructores no logran promover (Hestenes et al., 1992). Este estudio provee evidencia empírica de laboratorio a favor de maneras de fomentar estos procesos, empleando el potencial que ofrecen las diferencias de conocimiento que existen naturalmente en las aulas y encauzándolas a través de la elaboración por argumentación colaborativa de los conflictos sociocognitivos resultantes. Es decir, ofrece evidencia empírica con el potencial de nutrir investigaciones con mayor validez ecológica, basadas en el aula, sobre estrategias de instrucción en el paradigma de conflicto cognitivo para el cambio conceptual (Limón, 2001).

Finalmente, nuevas investigaciones podrían ayudar a traducir estos resultados en estrategias para manejar el crecimiento de ideas pseudocientíficas en general y sobre la forma de la Tierra en particular (Bär, 2018; Landrum, 2019; Moshakis, 2018; Parera, 2019; “The Challenge of the Post-Truth Era”, 2018). Reconocer la existencia de ideas equivocadas y las diferencias de conocimiento, y promover el conocimiento científico aprovechando la elaboración entre pares del conflicto sociocognitivo a través de la colaboración argumentativa, puede ayudar a allanar el camino que nos aleje del suelo infértil del negacionismo científico.

2.3 Estudio adicional

En este estudio evaluaremos un método alternativo para la descripción de modelos mentales basado en distancias entre conceptos en una red conceptual.[5]

2.3.1 Introducción

Los resultados del *Estudio principal 1* presentan evidencia a favor del uso de la enseñanza entre pares como estrategia para promover el aprendizaje mediado por procesos de cambio conceptual. Como discutimos al final, quedan preguntas sin resolver, que deberán abordarse en estudios futuros. Una de las limitaciones de

5 Los resultados de este estudio forman parte de un trabajo del curso de posgrado Seminario de Neurociencia Computacional, realizado por el tesista durante su doctorado.

estos estudios podría ser la dificultad en la ejecución de las entrevistas y su codificación, debido al tiempo que llevan y al riesgo de introducir sesgos de le experimentadore si no se controlan adecuadamente. Por lo tanto, en esta última parte del presente capítulo, describiremos el desarrollo de una posible alternativa para evaluar el conocimiento de les participantes, basada en la identificación de diferencias entre redes conceptuales mentales por la técnica del *análisis de la semántica latente* aplicada definiciones escritas.

Como desarrollamos en detalle en la Introducción del capítulo, durante el desarrollo ocurre un desplazamiento en la categorización del concepto "Tierra" hasta reconceptualizarlo como un objeto astronómico (Vosniadou et al., 2005, 2009). Es posible que estas modificaciones en el mapa de relaciones conceptuales de les sujetos se vean reflejadas en sus producciones escritas, evidenciando las diferencias entre redes conceptuales mentales pertenecientes a niñes de distintas edades y grados de desarrollo.

Una forma de explorar relaciones semánticas entre términos es mediante el uso de *word embeddings* (Altszyler et al., 2017); es decir, representaciones vectoriales de palabras o frases en un espacio semántico. En este trabajo intentamos identificar diferencias entre redes conceptuales mentales a través de la comparación de espacios semánticos obtenidos a partir de definiciones de distintos vocablos elaborados por niñes de distintas edades.

2.3.2 Métodos

2.3.2.1 Definiciones

Se utilizaron *definiciones* en catalán escritas por alrededor de 2400 niñes y adolescentes de entre 5 y 19 años (ver Figura 2.9). Éstas fueron obtenidas del sitio web del Proyecto CesCa (Llaurado et al., 2012), cuya finalidad, según les autores, es "proporcionar a la comunidad educativa una herramienta para conocer los usos lingüísticos de su alumnado". Es un *corpus* de referencia del catalán escolar escrito en Cataluña, e incluye textos de tipos diferentes, entre ellos definiciones de palabras.

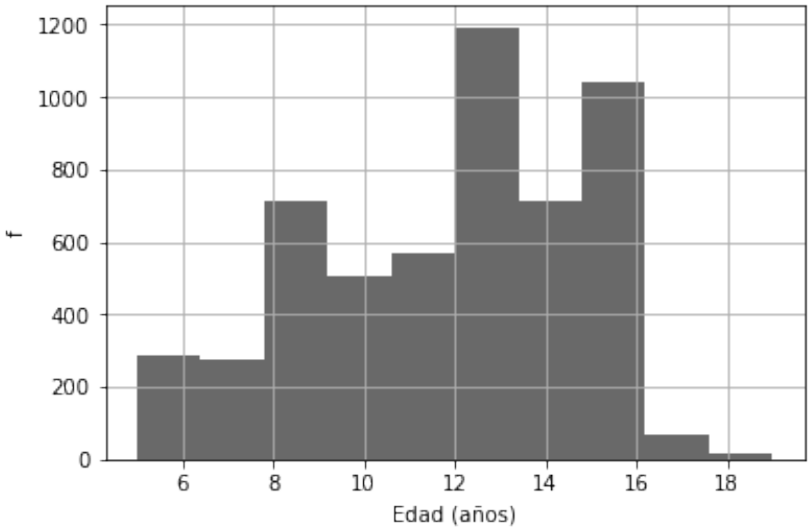


Figura 2.9: *Distribución del número de definiciones (f) por edad.*

Mediana = 12 años.

Si bien el corpus está libremente disponible para consulta en línea,[6] les autores no han previsto una vía para descargarlo. Por lo tanto, el primer paso del trabajo comprendió el desarrollo de un programa de consulta repetida al sitio y descarga automática del contenido. Éste fue escrito en lenguaje Python utilizando la librería *Splinter* para controlar automáticamente un navegador. Obtuvimos un total de 5381 definiciones útiles (es decir, correctamente etiquetadas por edad y lematizadas; ver a continuación), correspondientes a 66 vocablos diferentes (sustantivos, adjetivos y verbos). El código correspondiente y las definiciones así obtenidas están públicamente disponibles.[7]

Dado el tamaño relativamente pequeño del corpus, preferimos emplear una versión lematizada de las definiciones en los análisis subsiguientes, a fin de minimizar el tamaño del vocabulario o conjunto de términos (*tokens*) extraídos del corpus y hallar asociaciones más robustas. Es decir, una versión en la que palabras o conjuntos de palabras han sido reemplazados por su forma canónica o lema. Por ejemplo, el lema de las palabras “corre”, “corren”, “corrió” y “corriendo” es “correr”. Si bien existen formas automáticas de lematizar, por ejemplo empleando la librería FreeLing (Padró

6 http://clic.ub.edu/corpus/es/cesca-busquedas-corpus

7 https://github.com/diegodlh/cesca/blob/main/cesca/fetch_definitions.py, https://github.com/diegodlh/cesca/blob/main/cesca/definitions.csv

& Stanilovsky, 2012) con soporte para Catalán, preferimos usar la versión lematizada manualmente ofrecida por les autores, por considerarla más confiable.

Finalmente, para la construcción de los espacios semánticos, el vocablo definido se agregó al principio de la definición misma, para fortalecer la asociación entre éste y los términos empleados para definirlo (o sus lemas correspondientes).

2.3.2.2 Construcción de espacios semánticos

Primero, el corpus de definiciones se dividió en dos según la mediana de la edad de sus autores (ver Figura 2.9). Así, se obtuvo un subcorpus "menor" de 2352 definiciones correspondientes a les niños menores (de 5 a 11 años), y un subcorpus "mayor" de 3029 definiciones correspondientes a les niños mayores (de 12 a 19 años).

Luego, se definió un espacio semántico para cada subcorpus, aplicando la técnica de análisis de la semántica latente (*LSA: Latent Semantic Analysis)* a cada uno de ellos (Altszyler et al., 2017). Brevemente, primero (1) se construyó una matriz de co-ocurrencia término-definición para cada subcorpus. Es decir, una matriz dispersa (*sparse matrix*) que para cada definición indica el número de veces que aparece cada término del *vocabulario*. A continuación, (2) se aplicó una transformación (transformación Tf-idf: *term frequency-inverse document frequency*) con el objetivo de reducir el peso de términos poco informativos por su alta frecuencia en el corpus. Finalmente, (3) se hizo una reducción a 50 dimensiones (del orden de la cantidad de vocablos definidos: 66) por descomposición truncada en valores singulares (*truncated Singular-value decomposition*). El procedimiento paso a paso se encuentra públicamente disponible.[8]

8 https://github.com/diegodlh/cesca/blob/main/CesCa_corpus.ipynb.

2.3.2.3 Diferencias entre subcorpus

Para explorar las diferencias entre subcorpus, primero seleccionamos una serie de 40 lemas potencialmente relacionados con el modelo que les niñes tienen del planeta Tierra, algunos de los cuales estaban incluidos en el cuestionario y/o el cuadernillo del *Estudio principal 1* (ver secciones " Entrevista 1" y " Etapa de interacción entre pares o revisión autónoma"). Por ejemplo, "aigua" (agua), "aire" (aire), "astre" (astro), etc. La lista de lemas completa se encuentra en el repositorio en línea.[9]

Para evaluar la diferencia entre ambos grupos, evaluamos la distancia coseno (D), en uno y otro espacio semántico ("menor" y "mayor"), entre todos los pares de lemas posibles. De esta forma es posible responder para qué pares de lemas se da que su distancia (D) difiere entre el espacio definido para niñes menores (D_{men}) y aquel correspondiente a niñes/adolescentes mayores (D_{may}).

Para decidir si una diferencia de distancia ($\Delta D = |D_y - D_o|$) es significativa, se repitió el procedimiento anterior 1000 veces, pero permutando aleatoriamente cada vez la edad correspondiente a cada definición. Luego, para cada par de lemas, se calculó la media y el desvío estándar de la diferencia de distancia (ΔD). Así, se consideró que la distancia entre un par de lemas A y B cambió significativamente con el desarrollo (es decir, del grupo "menor" al grupo "mayor") si la diferencia de distancia (ΔD) obtenida es al menos cuatro desvíos estándar mayor que la media obtenida por permutación aleatoria. Nuevamente, el procedimiento en detalle puede consultarse en línea.[8]

2.3.3 Resultados

Efectivamente, la distancia entre algunos de los lemas seleccionados, relacionados con el tema de interés, difiere significativamente entre los subcorpus "menor" y "mayor" (Tabla 2.5).

9 https://github.com/diegodlh/cesca/blob/main/cesca/target_features.txt.

Lema 1	Lema 2	D_{men}	D_{may}	ΔD	$M_{\Delta D}$	$\sigma_{\Delta D}$
astre	constel·lació	0,429	0,008	-0,421	-0,019	0,096
constel·lació	espai	0,497	0,016	-0,481	-0,023	0,084
constel·lació	infinit	0,607	0,011	-0,597	-0,019	0,134
constel·lació	planeta	0,465	0,007	-0,458	-0,015	0,079
constel·lació	univers	0,620	0,011	-0,608	-0,019	0,130
envoltat	terra	0,264	0,026	-0,238	-0,007	0,039
illa	lluny	1,282	0,923	-0,360	0,000	0,064
illa	llunyà	1,288	0,924	-0,364	-0,001	0,064

Tabla 2.5: *Diferencias significativas de distancia entre lemas seleccionados*

D_{men}: distancia entre lemas en el subcorpus "menor". D_{may}: distancia entre lemas en el subcorpus "mayor". ΔD: diferencia entre distancias D_{may} - D_{men}; valores negativos indican convergencia entre lemas. $M_{\Delta D}$: diferencia de distancia media de las 1000 iteraciones de permutaciones de edad aleatoria. $\sigma_{\Delta D}$: ídem, desvío estándar. Se muestran sólo pares de lemas cuya diferencia de distancia es significativa:

$$|\Delta D| > M_{\Delta D} + 4 * \sigma_{\Delta D} \quad .$$

astre: astro; *constel·lació*: constelación; *espai*: espacio; *infinit*: infinito; *planeta*: planeta; *univers*: universo; *evoltat*: rodeado; *terra*: tierra; *illa*: isla; *lluny*: lejos; *llunyà*: lejano.

Resulta de particular interés el caso del término "constel·lació" (constelación). Este término aparece en más definiciones del grupo "mayor" que del grupo "menor". De las 11 definiciones del grupo "mayor" en las que aparece "constel·lació", 8 (73%) son definiciones del término "univers" (universo). En el grupo "menor", en cambio, de las sólo 5 definiciones en las que aparece "constel·lació", sólo una corresponde al vocablo "univers". Esto a pesar de que el número de definiciones totales del vocablo "univers" es similar en ambos grupos (36 para el grupo "menor" y 41 para el grupo "mayor", respectivamente).

2.3.4 Discusión

El objetivo de este estudio era evaluar la utilidad de la técnica del análisis de la semántica latente (LSA) para simplificar la evaluación y comparación del conocimiento de les participantes en futuros estudios, como estrategia alternativa a las entrevistas y la codificación empleadas en el Estudio principal.

Concretamente, pudimos identificar diferencias de distancia entre conceptos (expresados como términos o lemas) entre dos corpus derivados de definiciones producidas por niñes entre 5 y 11 años, por un lado, y niñes/adolescentes de 12 a 19 años, por otro. Es posible que estas diferencias reflejen cambios en las redes conceptuales mentales que ocurren durante el desarrollo en el dominio de conocimiento considerado, como fue mencionado en la Introducción de este capítulo.

En particular, observamos que en les participantes mayores el concepto “constelación” se encontraría más cerca de los conceptos “astro”, “espacio”, “infinito”, “planeta” y “universo”. Un análisis más detallado muestra que esto podría deberse a que les individues de menor edad usaron menos la palabra “constelación” para definir el vocablo “universo”, y ésta podría ser la razón por la que “constelación” y “universo” (y otros conceptos relacionados: “astro”, “espacio”, “infinito”, “planeta) se encuentran más próximos en el espacio semántico correspondiente a les individues mayores.

Cabe mencionar que es probable que la metodología haya estado limitada por la reducida longitud de los documentos (definiciones) que forman parte del corpus. Partir de producciones escritas más extensas podría sortear este inconveniente, aunque restringiría la edad de les participantes, puesto que niñes más pequeñes podrían encontrar dificultades para producir textos muy largos. Alternativamente, se podría trabajar con transcripciones de producciones orales, aunque son más difíciles de obtener.

En resumen, los resultados obtenidos sugieren que esta metodología podría utilizarse para simplificar la evaluación y comparación de los modelos mentales de les participantes en estudios futuros que busquen profundizar los resultados obtenidos en el Estudio principal de este capítulo.

2.4 Conclusiones

En este capítulo abordamos la cuestión del cambio conceptual colectivo e individual promovido por la interacción social, en particular por el conflicto sociocognitivo, posiblemente resuelto a través de argumentación colaborativa mediada por la enseñanza entre pares, el tema general de esta tesis.

Nuestros resultados sugieren que la interacción entre niñes con modelos mentales diferentes en el campo de la astronomía observacional promueve procesos de cambio conceptual que resultan en la reducción de la brecha de conocimiento entre elles. En primer lugar, la evaluación de niveles de conocimiento mostró que esta reducción no fue el resultado de una convergencia hacia acuerdos intermedios, sino que fue mediada por procesos de cambio conceptual en les niñes menos avanzades. En segundo lugar, la evaluación de la similitud de conocimiento mostró que este avance no sería el producto de construcciones individuales independientes, sino de la construcción colectiva de conocimiento mediada por la adopción de respuestas de les participantes más avanzades. Finalmente, la evaluación a la luz de la teoría sobre el cambio conceptual en este dominio es compatible con cambios hacia modelos no duales, sugiriendo que la interacción podría haber catalizado la revisión laboriosa de presuposiciones arraigadas tempranamente, acelerando estos procesos de cambio conceptual y reconceptualización ontológica.

Cabe resaltar la naturaleza cuantitativa de los métodos de análisis empleados. En este sentido, en el Estudio adicional presentamos una posible alternativa para la evaluación y comparación del conocimiento de les participantes, basada en distancias entre conceptos en un mapa conceptual construido a partir de sus elaboraciones escritas, que podrían permitir acelerar la recolección y análisis de aún más datos en investigaciones futuras.

En resumen, nuestros resultados aportan evidencia para el uso de la enseñanza entre pares como estrategia pedagógica para promover procesos de cambio conceptual, y abren preguntas para abordar en nuevos estudios, proponiendo a su vez un método alternativo de análisis del conocimiento de les participantes. Sin embargo, algunos aspectos relevantes de la tesis no fuero abordados en este capítulo. Por un lado, el tipo de tutoría entre pares empleada en este capítulo no es típica. ¿Qué impacto tendría sobre el aprendizaje una situación en la que la dirección del flujo de información fuese más marcada? Tampoco fue abordada la relación con otras habilidades cognitivas, como la metacognición, con el desarrollo, o con otros dominios del conocimiento. ¿Qué efecto tendría sobre procesos de aprendizaje no ya de cambio conceptual, pero sí igualmente dificultosos e importantes para el mundo en el que vivimos? Ni agotamos la cuestión sobre el posible efecto de enseñar, del que no encontramos evidencia en este capítulo;

aunque sí identificamos posibles causas de este resultado, que entendimos importantes considerar en los estudios siguientes.

En el siguiente capítulo, en el que tratamos la cuestión de la enseñanza entre pares para el caso específico del aprendizaje de lenguajes de programación, son abordados algunos de estos aspectos.

Capítulo 3: Enseñanza entre pares de lenguajes de programación

En el Capítulo 2 abordamos la cuestión de la enseñanza entre pares como promotora del cambio conceptual en niñes. Sin embargo, una de las limitaciones del paradigma de enseñanza entre pares empleado era que la asimetría de roles y conocimiento no era explícita.

En este capítulo empleamos un dominio del conocimiento en el que les niñes carecen de conocimientos previos, permitiendo así el establecimiento de una asimetría clara entre tutore y aprendiz: el aprendizaje de lenguajes de programación.

Además, abordamos la cuestión de la interacción entre la enseñanza entre pares y otras habilidades cognitivas. En particular, con la metacognición.

3.1 Introducción

El mundo de hoy está inundado por tecnologías digitales. Las computadoras nos rodean, y se comunican entre ellas de forma permanente. Esta red creciente de computadoras se encuentra profundamente entretejida con nuestra red social y subyace muchos aspectos de nuestras sociedades actuales (Dijk, 2010; Fundación Sadosky, 2016; President's Council of Advisors on Science and Technology, 2010). Con las computadoras y sus redes en los cimientos de nuestra sociedad mundial, entender cómo operan y cómo moldearlas nos permite progresar de un rol de usuarios pasivos a uno de creadores activos (Resnick et al., 2009), y es fundamental para nuestra libertad (Söderberg, 2012).

La programación de computadoras es la forma que nosotres, les humanes, tenemos para comunicarnos con estas máquinas, y las redes a las que éstas pertenecen, y así controlar sus funciones. Así, programar consiste en diseñar e implementar conjuntos de instrucciones interpretadas por una computadora para cumplir una tarea específica (Dalbey & Linn, 1985; Fundación Sadosky, 2016; Papert, 1980).

Existe una gran variedad de lenguajes de programación, y comprenderlos se está volviendo cada vez más importante (Fedorenko et al., 2019; Fundación Sadosky, 2016; OPSSI, 2019). Además, aprender a programar no sólo sirve para adquirir

estas habilidades de programación, sino que también hay evidencia de que promueve el desarrollo de otras habilidades cognitivas, tales como la creatividad, habilidades de cognición espacial, habilidades matemáticas, razonamiento lógico, entre otras (Fundación Sadosky, 2016; Papert, 1980; Resnick et al., 2009; Scherer, 2016; Scherer et al., 2018), algunas de las cuales se conocen colectivamente como *Pensamiento Computacional* (Grover & Pea, 2013; Wing, 2017).

Les niñes se vuelven usuaries de *software* y *hardware* a edades cada vez más tempranas (Rideout, 2013). Sin embargo, ser une native digital no viene necesariamente acompañado de entender cómo funcionan estas tecnologías (Martinez et al., 2015). Por esta razón, existe un interés creciente por integrar cursos de programación en todos los niveles educativos (Fessakis et al., 2013; Gülbahar & Kalelioğlu, 2014). Muchas escuelas, distritos y países han reconocido la importancia de incluir la programación en sus programas de estudio (Balanskat & Engelhardt, 2015; Brinda et al., 2009; Brown et al., 2013; Consejo Federal de Educación de Argentina, 2018; Fundación Sadosky, 2016; Kalelioğlu, 2015), y algunes incluso la consideran una nueva forma de alfabetización, como la lectura o la matemática (Fedorenko et al., 2019; O′Reilly, 1998; Voogt et al., 2015; Wing, 2006). En Argentina, por ejemplo, fue declarada “de importancia estratégica para el sistema educativo (…) durante la escolaridad obligatoria, para fortalecer el desarrollo económico-social de la Nación” (Consejo Federal de Educación de Argentina, 2015).

Pero enseñar a les niñes esta capacidad no es tarea sencilla. Recientemente, han surgido nuevos lenguajes de programación que creativamente han conseguido reducir ciertas dificultades sintácticas y semánticas de los lenguajes de programación tradicionales (Wyeth, 2008), al punto de ser totalmente visuales. Estos lenguajes visuales permiten a les usuaries manipular sus elementos gráficos en vez de especificarlos textualmente, y por lo tanto resultan particularmente adecuados para niñes (Maloney et al., 2010; Resnick et al., 2009). Sin embargo, enseñar a programar presenta algunas dificultades características: (i) la variabilidad de métodos de enseñanza, sin un claro entendimiento acerca de cuáles son más adecuados en qué contextos; (ii) la falta de estándares de enseñanza y de docentes capacitados (Fundación Sadosky, 2016); (iii) una larga tradición en las clases de Tecnologías de la Información y la Comunicación de enfocarse en habilidades de computación denominadas *blandas*, como el uso de procesadores de textos,

planillas de cálculos, y otras herramientas de oficina u ofimáticas (Benotti et al., 2014; Fundación Sadosky, 2016; Resnick et al., 2009); y (iv) diferencias motivacionales y de desempeño entre géneros (Edwards et al., 1997; Hill et al., 2010; Nourbakhsh et al., 2004), muchas de ellas producto de convenciones y estereotipos sociales (Beilock et al., 2010; Dar-Nimrod & Heine, 2006); entre otras.

Algunas estrategias adoptadas por la industria del software para optimizar sus procesos podrían echar luz sobre esta problemática. La programación de a pares, por ejemplo, es una práctica que involucra a dos desarrolladores colaborando como une sole individue en el diseño, escritura y prueba de la misma tarea de programación. Esta práctica ha mostrado ser productiva y generar código de mayor calidad que el que cualquiera de les desarrolladores podría generar por su cuenta (Lui & Chan, 2006; Nosek, 1998; Williams, 2001; Williams et al., 2000; Williams & Upchurch, 2001). Esta observación sería particularmente aplicable a programadores novates, y a problemas de programación desafiantes (Lui & Chan, 2006). Además, la programación entre pares también ha mostrado efectos positivos para el aprendizaje en contextos educativos (Braught et al., 2011; Umapathy & Ritzhaupt, 2017).

Otra práctica popular es la de tomar una porción de código problemática, y explicar línea por línea lo que supuestamente debería hacer. Esto fuerza a le programadore a clarificar y estructurar la explicación, ayudándole a depurar el código, es decir a encontrar la causa del problema. Esta estrategia funciona incluso cuando el código se explica a un objeto inanimado, razón por la cual se la suele conocer como depuración con patito de goma o *rubber duck debugging* (Hunt & Thomas, 1999; Phillips et al., 2013).

Dados el creciente interés en la enseñanza temprana de programación, las experiencias de la industria del software, y la efectividad reconocida de la enseñanza entre pares, es sorprendente que no se hayan hecho estudios experimentales aún sobre enseñanza de programación entre pares en niñes y cómo podría ayudar a resolver algunas de las dificultades de enseñar a programar. En especial considerando que les niñes son natives digitales y que a veces incluso superan a sus docentes en estos temas (Benotti et al., 2014). Algunos estudios se han aproximado a la temática, pero ninguno la ha hecho el foco de su atención. Por ejemplo, algunos de ellos han considerado la adquisición de competencias básicas de computación por parte de grupos de niñes sin supervisión, pero no incluyeron

habilidades de programación (Mitra et al., 2005). Por otra parte, hay estudios sobre enseñanza entre pares y evaluación entre pares de programación, pero se han enfocado en estudiantes de edad más avanzada, en escuelas secundarias y universidades (Altintas et al., 2016; Golding et al., 2006), y algunos de ellos son de tipo exploratorio y no experimentales (Edwards et al., 1997).

Por lo tanto, en este estudio exploramos la enseñanza entre pares como una forma de ayudar a les niñes a aprender lenguajes de programación. Para este fin adaptamos una plataforma de programación comercial para niñes. Primero, estudiamos si les niñes pueden enseñar a otres niñes a programar comparando las habilidades de programación y estrategias resolutivas desarrolladas por (i) niñes que aprendieron de une adulte, vs (ii) niñes que aprendieron de une par. Por otra parte, también analizamos si les niñes obtienen un beneficio de enseñar a programar, comparando (i) niñes que enseñaron a une par, vs (ii) niñes que revisaron el material a solas. Las diferencias entre géneros, la relación con la metacognición y el impacto a largo plazo de las distintas condiciones fueron evaluados también.

3.2 Estudio principal 2

En este estudio evaluaremos la tutoría entre pares como estrategia pedagógica para la enseñanza de lenguajes de programación.[10]

3.2.1 Métodos

3.2.1.1 Muestra

44 estudiantes de dos segundos grados, de una escuela primaria bilingüe de nivel socioeconómico medio-alto del Área Metropolitana de Buenos Aires, participaron del estudio (16 de género masculino y 28 de género femenino, según fue informado por las autoridades de la escuela). Los datos de dos niñas fueron descartados, en un caso porque su sesión de evaluación (ver sección “3.2.1.5 Procedimiento”) tuvo que ser interrumpida por razones de fuerza mayor, y en el otro porque la participante tenía dificultades en comprender el español. Las edades de les niñes restantes

10 Los resultados de este estudio se encuentran en vías de publicación en: de la Hera, D.P., Zanoni, M. B., Sigman, M., & Calero, C. I. (en prensa). Peer tutoring of computer programming increases exploratory behavior in children. *Journal of Experimental Child Psychology*.

variaron de 7 años y 1 mes, a 8 años y 4 meses, con una media de 7 años y 10 meses. Todas las sesiones tuvieron lugar en una sala tranquila, provista por la escuela para los propósitos del estudio, durante noviembre de 2016.

Les padres o tutores de todes les participantes firmaron voluntariamente el consentimiento previamente autorizado por el Comité de Ética de la Dirección de Investigación del Centro de Educación Médica e Investigación Clínica “Norberto Quirno” (CEMIC), Unidad Asociada del CONICET, Protocolo Nº 683.

3.2.1.2 Plataforma de programación

Utilizamos la plataforma comercial WeDo 1 de LEGO Education[11] (Mayerová & Veselovská, 2014), comercializada para niñes a partir de los siete años (Kazakoff et al., 2013), y usada previamente en otros estudios con niñes alrededor de esta edad (Mayerové & Veselovská, 2017; Pinto-Llorente et al., 2018). En el contexto de esta tesis, entenderemos a la plataforma de programación como comprendida por (i) un lenguaje de programación, (ii) un editor de programas, (iii) un intérprete, y (iv) un entorno programable. El lenguaje de programación (i) es un lenguaje visual, basado en íconos, similar a *Scratch Jr* (Flannery et al., 2013). Comprende una serie de bloques, cada uno expresando un comando diferente. En el editor de programas (ii), los bloques se toman del panel de bloques y se arrastran hasta el lienzo del editor, donde se acoplan uno a continuación del otro para construir secuencias de bloques o programas (Figura 3.1A). El intérprete (iii), por su parte, lee estos programas y opera sobre el entorno programable (iv), alterando su configuración o estado (Turing, 1936). El entorno programable incluye (a) una pantalla virtual, ubicada en el extremo superior izquierdo de la pantalla, (b) los parlantes de la computadora, y (c) un robot construido con bloques LEGO con un brazo giratorio, asemejando un reloj mirando hacia arriba.

11 LEGO y la minifigura son marcas comerciales y propiedades intelectuales de The LEGO Group. ©2021 The LEGO Group.

Figura 3.1: *Métodos*

(A) Captura de pantalla. La figura muestra una captura de pantalla de la plataforma de programación, con una secuencia de bloques construida en el lienzo del editor de programas, y la interfaz de usuarie en modo *evaluación*, en el ejercicio 2, y con el video del ejercicio pausado en el segundo 3. El lienzo del editor fue comprimido en esta captura para optimizar la presentación. (1) Panel de bloques del editor de programas. (2) Lienzo del editor de programas. (3) Pantalla del entorno programable, con img=11 y txt="" (cadena de texto vacía). (4) Reproductor de videos en pausa; el número de ejercicio actual se muestra en la parte superior ("JUEGO 2"). (5) Botón de reproducción/pausa, exhibido como reproducción porque el video se encuentra pausado. (6) Botón de recarga. (7) Botón "ejercicio siguiente". (8) Botón "ejercicio anterior". (9) Botón "reiniciar entorno programable".

(B) Ejemplo de video. *Arriba*: se muestran cuatro cuadros de un video de ejemplo. El tiempo de reproducción se muestra en la esquina inferior izquierda de cada cuadro. Cada cuadro representa un estado diferente del entorno programable, descripto por el valor de cinco variables, mostradas aquí en un recuadro en la esquina inferior derecha de cada cuadro (img: imagen, txt: texto, pos: posición, rot: rotación, snd: sonido; ver texto principal). Ni el tiempo de reproducción ni el

recuadro de variables aparecen en los videos reales. *Abajo*: debajo de los cuadros del video se muestra la secuencia de bloques o programa que recrea la serie de eventos representada en el video. (la ilustración del pato fue diseñada por brgfx / Freepik).

El lenguaje de programación se puede entender como operando sobre cinco variables diferentes que describen el *estado* del entorno programable (Turing, 1936). (1) La variable *imagen* (img) describe la imagen que se muestra en la pantalla virtual del entorno, y puede tomar valores numéricos del 1 al 20, cada uno indicando una imagen diferente, o 0 para indicar una imagen en blanco. (2) La variable *texto* (txt) describe el texto que se muestra en la pantalla del entorno, y puede tomar como valor cualquier cadena de texto indicada por le usuarie. (3) La variable *posición* (pos) describe la posición del brazo robótico en grados, y puede tomar valores numéricos entre 0 y 360. (4) La variable *rotación* (rot) describe si el brazo robótico se está moviendo y en qué sentido, y puede tomar los valores +1, -1 o 0, indicando rotación horaria, antihoraria, o sin rotación, respectivamente. Por último, (5) la variable *sonido* (snd) describe el sonido que está siendo reproducido por los parlantes de la computadora, y puede tomar valores numéricos del 1 al 20, cada uno indicando un sonido diferente, y 0 para ningún sonido o silencio. Les usuaries tenían a disposición un catálogo impreso a dos caras, incluyendo todas las imágenes disponibles y sus valores correspondientes de un lado, y todos los sonidos y sus valores del otro (ver “Anexo 3.1: Catálogo de imágenes y sonidos”).

Al comienzo de cada ejercicio, el entorno programable se encuentra en el estado inicial: [img=0, txt=“” (cadena vacía), pos=0, rot=0, snd=0]. Las secuencias de bloques o programas se construyen en el lienzo del editor de programas acoplando *bloques de comando* uno a continuación del otro.

El intérprete lee los programas de izquierda a derecha y actualiza el estado del entorno programable según indique el programa (Figura 3.1B), como una máquina de Turing (Turing, 1936). Por ejemplo, un bloque de comando “imagen” (ver apartado *Tipos de bloques*, a continuación) indica al intérprete que actualice la variable img del entorno programable. Nos referiremos a estos cambios en el estado del entorno programable como *eventos.*

Tipos de bloques

En la adaptación de la plataforma realizada para este estudio, existen ocho *bloques de comando* diferentes. Algunos de ellos (*transitivos*) admiten un *bloque de argumento* adicional acoplado abajo para definir el comando con mayor precisión. Además, los bloques se dividen a su vez en tres tipos:

(i) Bloques de control: controlan el comportamiento del intérprete, sin actualizar ninguna variable. *Bloque de inicio:* bloque cliqueable a partir del cual el intérprete comienza a leer una secuencia de bloques. Éste es el único bloque que no admite bloques acoplados a su izquierda. Cuando se hace clic en un bloque de inicio, el intérprete lee el bloque acoplado a su derecha (si es que hay uno) y continúa con el siguiente, hasta el último bloque en la secuencia. *Bloque de espera:* retiene al intérprete la cantidad de tiempo especificada en el bloque de argumento acoplado abajo (en décimas de segundo), antes de continuar con el bloque siguiente.

(ii) Bloques de acción: el intérprete los lee, actualiza la variable correspondiente, y procede inmediatamente con el bloque siguiente. *Bloques de imagen y de texto:* actualizan la variable img o txt, respectivamente, según el valor especificado en el bloque de argumento acoplado abajo. *Bloques rotar horario, rotar antihorario, y detener:* cambia la variable rot a +1, -1 o 0, respectivamente.

(iii) Bloques de acción y control: el intérprete los lee, actualiza la variable correspondiente, y es retenido durante un tiempo predefinido antes de continuar con el bloque siguiente. *Bloque de sonido:* cambia la variable snd al valor especificado en el bloque de argumento acoplado abajo, espera el tiempo que dure el sonido seleccionado, y finalmente cambia la variable snd nuevamente a 0 (silencio).

La variable pos (posición) es la única que no es actualizada explícitamente con un bloque de comando, sino que en cambio es actualizada continuamente según el valor de la variable rot. En otras palabras, les usuaries no pueden ajustar explícitamente la variable pos a 180, por ejemplo, sino que deben usar el bloque “rotar horario” para ajustar la variable rot a +1 (rotación horaria), luego usar un bloque “de espera” para indicar al intérprete que espere una cierta cantidad de tiempo para proseguir, mientras la variable pos aumenta hasta 180, y finalmente usar el bloque “detener” para ajustar la variable rot nuevamente a 0 (Figura 3.1B).

3.2.1.3 Interfaz de usuarie

Para este estudio, desarrollamos una interfaz de software interpuesta entre le usuarie y la plataforma de programación WeDo. El objetivo fue evitar interacciones no deseadas con la plataforma, y proporcionar un marco adecuado que incluyera una serie de ejercicios para evaluar el nivel de aprendizaje alcanzado por les participantes (Figura 3.1A). Esta interfaz fue programada con TkInter,[12] una adaptación para Python de las herramientas Tk de interfaz gráfica de usuario, y el código fuente está disponible públicamente.[13]

Control de interacciones

Con respecto al control de la interacción del usuario con la plataforma, para simplificar el lenguaje de programación redujimos el número de bloques disponibles en el panel de bloques del editor de programas a un subconjunto limitado de ocho, ocultando y deshabilitando los demás. Por ejemplo, excluimos bloques que sirvieran de atajo, para promover el aprendizaje de los conceptos de programación básicos; como el bloque “rotar durante”, que es un atajo de la secuencia rotar+esperar+detener. O bloques no relacionados, como los bloques de operación matemática. O bloques avanzados, como el bloque “repetir”, o los bloques de entrada (*input*).

Además, otras interacciones no deseadas con la plataforma y con el sistema operativo fueron asimismo bloqueadas, como por ejemplo la posibilidad de ocultar el panel de bloques o la pantalla virtual del entorno programable, de cerrar la plataforma de programación, de apagar la computadora o ajustar el volumen, y una larga lista de otras posibilidades.

Reproductor de videos y navegador de ejercicios

Por otra parte, a fin de proporcionar un marco adecuado para la evaluación, incluimos un reproductor de videos y un navegador de ejercicios (Figura 3.1A). El reproductor de videos mostraba un video por ejercicio, cada uno presentando una serie de eventos (es decir, de cambios en el estado del entorno programable) que les participantes debían reproducir construyendo sus propios programas (ver Figura

12 https://wiki.python.org/moin/TkInter

13 http://gitlab.com/diegodlh/lego-wedo

3.1B y la sección “3.2.1.4 Videos”*,* más adelante). El reproductor de video contaba con un botón de reproducción/pausa para iniciar, pausar y reanudar la reproducción tantas veces como se deseara, y un botón de recarga para reiniciar la reproducción desde el comienzo. El navegador de ejercicios, por su parte, incluía botones “anterior” y “siguiente” para retroceder y avanzar a través de los ejercicios a gusto. Cada vez que un ejercicio era accedido con el navegador de ejercicios, se contaba como un *intento* para dicho ejercicio, y la posición de los bloques en el lienzo del editor era recordada entre intentos del mismo ejercicio. Finalmente, les participantes podían reiniciar el estado del entorno programable en cualquier momento usando un botón de reinicio dispuesto para tal fin (Figura 3.1A). Cabe destacar que esta función de reinicio no es soportada de forma nativa para el brazo robótico, por lo que tuvimos que desarrollar una solución específica empleando sensores ópticos y controladores de *software* alternativos.[14]

Pantalla de autoevaluación

Como les participantes podían avanzar al ejercicio siguiente en cualquier momento, sin importar su desempeño en el ejercicio en curso, se introdujo una instancia de autoevaluación entre ejercicio y ejercicio, para desalentar el uso compulsivo del botón “siguiente” en el navegador de ejercicios, y para medir la confianza de les participantes en relación con su desempeño. Esta instancia de autoevaluación se mostraba al pulsar el botón “siguiente” al concluir el primer intento de cada ejercicio. En ella se preguntaba a le participante si había podido copiar el video, y se ofrecían tres opciones de respuesta, acompañadas de una imagen representativa de un semáforo: “sí” (luz verde), “más o menos” (luz amarilla), “no” (luz roja). Si le usuarie respondía “sí”, la autoevaluación finalizaba y se avanzaba al ejercicio siguiente. En cambio, si le usuarie respondía “más o menos” o “no”, se le preguntaba si deseaba continuar con el ejercicio siguiente, o regresar al ejercicio actual.

Modos de funcionamiento

La interfaz de usuarie dedicada descripta en esta sección podía ser ejecutada en uno de tres modos posibles: *introducción, práctica, y evaluación/reevaluación.* El modo *introducción* no incluía el reproductor de video ni el navegador de ejercicios. El modo *práctica* incluía el reproductor de videos y el navegador de ejercicios con dos

14 https://github.com/itdaniher/WeDoMore

ejercicios de ejemplo, pero sin autoevaluación. Los modos *introducción* y *práctica* se usaron en la primera y en la segunda parte, respectivamente, de las etapas de *Instrucción* y de *Enseñanza entre pares/Revisión* (ver sección "3.2.1.5 Procedimiento", más adelante). Finalmente, el modo *evaluación*/*reevaluación* fue usado en las etapas homónimas. Incluía el reproductor de videos y el navegador de ejercicios con diez ejercicios de evaluación o reevaluación, respectivamente, más un ejercicio de prueba adicional al comienzo (ejercicio 0), con un video sin eventos para explicar a les participantes cómo usar el reproductor de videos, el navegador de ejercicios, y la autoevaluación (ver sección "3.2.1.5 Procedimiento"). En este modo, un contador se iniciaba al comienzo del ejercicio 1; este contador congelaba la pantalla, impidiendo todo tipo de interacción, después de 20 minutos.

3.2.1.4 Videos

Los videos de los ejercicios muestran una serie de *eventos*, cada uno involucrando una o más de las variables `img`, `txt`, `pos`, `rot` y `snd`, que les participantes deben reproducir ensamblando secuencias de bloques o programas (Figura 3.1B). Estos videos incluyen: (i) una copia de la pantalla virtual del entorno programable en la esquina superior izquierda, para indicar cambios en las variables `img` y `txt`; (ii) un área en la esquina superior derecha donde aparecen íconos representando los distintos sonidos posibles, acompañados del sonido correspondiente, para indicar cambios en la variable `snd`; y (iii) una representación del robot y de su brazo rotatorio visto desde arriba, para indicar cambios en las variables `pos` y `rot`. Todos los videos comienzan representando el entorno programable en su estado inicial [`img`=0, `txt`="", `pos`=0, `rot`=0, `snd`=0]; el primer evento ocurre al segundo 1 y los demás cada 2 segundos (ver Figura 3.1B para un ejemplo; todos los videos están disponibles públicamente).[15]

En las etapas de *Evaluación* y de *Reevaluación*, se presentaron 10 videos mostrando series de eventos de complejidad creciente (ver "Anexo 3.4: Eventos en videos"). Los videos fueron presentados en el mismo orden para todes les participantes. Los videos en la reevaluación eran isomórficos con los videos en la evaluación, es decir que mostraban eventos involucrando a las mismas variables, pero con valores diferentes para hacerlos más interesantes (Smith et al., 2009). Por

15 https://gitlab.com/diegodlh/lego-wedo/tree/master/videos

ejemplo, si el segundo evento en el video nro. 3 de la etapa de *Evaluación* implicaba cambiar la variable `img` de 0 a 5 (imagen del bosque), entonces el segundo evento del video nro. 3 de la etapa de *Reevaluación* podría implicar cambiar la variable `img` también, pero de 0 a 7 (imagen del desierto).

3.2.1.5 Procedimiento

El procedimiento se dividió en cinco etapas: (I) de Instrucción; (II) de Evaluación inicial; (III) intermedias, de Enseñanza entre pares o de Revisión autónoma; (IV) de Evaluación final, y (V) de Seguimiento (ver Figura 3.2).

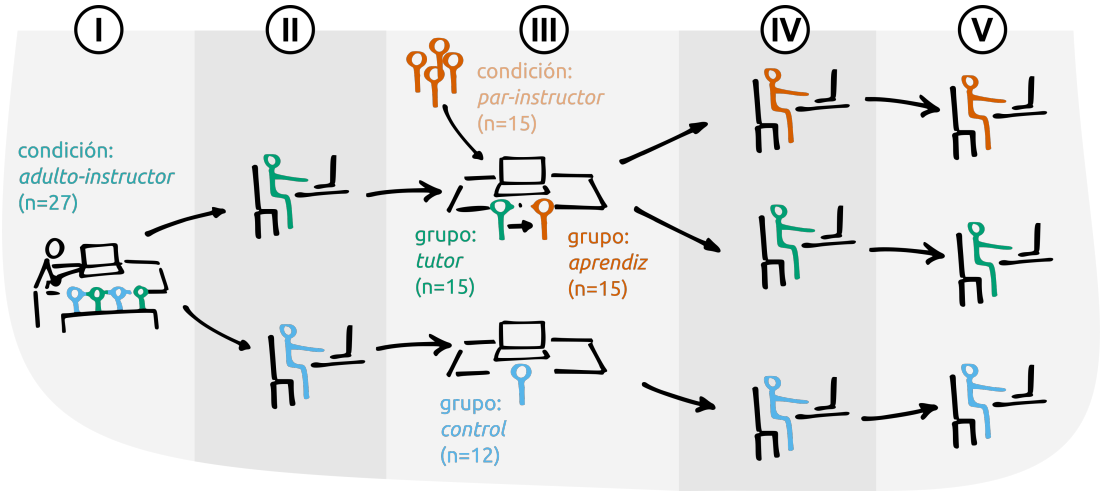


Figura 3.2: *Procedimiento*

(I) Etapa de Instrucción. Les niñes en la condición *adulto-instructor* reciben una clase a cargo de une instructore adulte.

(II) Etapa de Evaluación inicial. Les niñes en la condición *adulto-instructor* son evaluades por primera vez (Etapa de Evaluación).

(III) Etapas intermedias. Les niñes en la condición *adulto-instructor* son divides en dos grupos: *tutor* y *control*. *Arriba: etapa de Enseñanza entre pares.* Les niñes del grupo *tutor* enseñan a compañeres del grupo *aprendiz* (condición *par-instructor*) que no participaron de las etapas anteriores. *Abajo: etapa de Revisión autónoma.* Les niñes del grupo *control* revisan el material a solas.

(IV) Etapas de evaluación final. Les niñes de los grupos *tutor* y *control* (condición *adulto-instructor*) son evaluades por segunda vez (Etapa de Reevaluación). Les niñes del grupo *aprendiz* (condición *par-instructor*) son evaluades por primera vez (Etapa de Evaluación).

(V) Etapa de Seguimiento. Todes les participantes son evaluades nuevamente dos años más tarde.

Etapa de Instrucción

Primero, dos terceras partes de les participantes en cada clase fueron asignades a la condición *adulto-instructor* (n=27; el resto fue asignado a la condición *par-instructor* –ver más adelante). En esta condición, grupos de cuatro a cinco niñes del mismo curso participaron de una clase de 20 minutos estrictamente guionada (ver "Anexo 3.5: Guión de instrucción") a cargo de une instructore adulte[16]. Les niñes se sentaban une al lado de le otre, frente a la pantalla de una computadora portátil, con el instructor de pie a su izquierda. Las clases comprendían dos partes.

En la primera parte, se enseñaban los fundamentos del lenguaje de programación, siguiendo un libro de instrucciones de 23 páginas (ver "Anexo 3.2: Libro de instrucciones") y usando la interfaz de usuarie en el modo *introducción*, sin reproductor de video ni navegador de ejercicios. El robot permanecía oculto para evitar distracciones, y era revelado por el instructor en el momento específico indicado por el guión.

En la segunda parte, se presentaban dos ejercicios de ejemplo, siguiendo un libro de ejemplos (ver "Anexo 3.3: Libro de ejemplos") y utilizando la interfaz de usuarie en modo *práctica.* En esta parte, le instructore adulte pausaba los videos de ejemplo en cada evento y construía la secuencia de bloques que permitía su reproducción. Además, el instructor presentaba y empleaba el catálogo impreso a dos caras, incluyendo todas las imágenes disponibles y sus valores correspondientes de un lado, y todos los sonidos y sus valores del otro (ver "Anexo 3.1: Catálogo de imágenes y sonidos"). Al final de cada ejemplo, se mostraba la solución correcta en el libro de ejemplos y se la comparaba con la secuencia construida.

En ambas partes, tanto el *mouse*, como el teclado y los materiales de instrucción estaban bajo control exclusivo de le instructore adulte.

Etapa de Evaluación inicial

Luego, en promedio 4 días más tarde (mín=2, máx=7) les niñes enseñades por une adulte fueron evaluades individualmente en 10 ejercicios, con la interfaz de usuarie

16 Todas las clases estuvieron a cargo de le misme instructore adulte, de género masculino.

en el modo *evaluación*. Siguiendo estrictamente un guión (ver “Anexo 3.6: Guión de explicación”), se le indicaba a le participante que debía usar los bloques del panel de bloques del editor de programas para reproducir los eventos que veía en los videos. También se explicaban el reproductor de videos, el navegador de ejercicios, y los botones de la instancia de autoevaluación. A continuación, el investigador entregaba a le participante el catálogo de imágenes y sonidos (ver “Anexo 3.1: Catálogo de imágenes y sonidos”), iniciaba el ejercicio 1, y abandonaba la sala. En caso que le participante saliera de la sala para hacer preguntas, el investigador le acompañaba nuevamente al interior y se limitaba a responderle con un resumen predefinido de la explicación original.

Al finalizar la sesión, 20 minutos después, el investigador ingresaba a la sala nuevamente. Intencionalmente, este tiempo no era suficiente para resolver los 10 ejercicios correctamente. Esto permitía reservar cierto margen de mejora para la etapa de *Reevaluación*, con el objetivo de evitar efectos techo (que podrían haber ocultado un posible *efecto tutor* en el Estudio principal del Capítulo 2). El investigador simulaba revisar rápidamente las respuestas de le participante, felicitándole por su trabajo en cada ejercicio, pero sin hacer devoluciones adicionales acerca de su desempeño.

Etapas intermedias

A continuación, les niñes enseñades por une adulte (condición *adulto-instructor*) fueron dividides en dos grupos de desempeño equivalente: *tutor* (n=15) y *control* (n=12). Un tercer grupo *aprendiz* (n=15) incluía a les niñes que no habían participado de las etapas previas de Instrucción y Evaluación (condición *par-instructor*).

Etapa de Enseñanza entre pares

Les niñes en el grupo *tutor* fueron emparejades con compañeres del mismo género y del mismo curso del grupo *aprendiz*, y se les pidió que enseñaran a su compañere para que éste pudiera jugar también (en promedio 7 días después de la etapa de *Evaluación*). Brevemente, ambes participantes se sentaban del mismo lado de una mesa, de frente a la pantalla de la computadora, con le tutore sentade a la izquierda de su aprendiz, y el robot fuera de vista, como en la etapa de *Instrucción*.

En la primera parte, la interfaz de usuarie era ejecutada en modo *introducción*. En frente de le aprendiz, le tutore era felicitade por su desempeño en la etapa de *Evaluación* y se le pedía que enseñara a su compañere para que pudiera jugar también. Le tutore recibía el libro de instrucciones y el investigador abandonaba la sala. Cuando fuera llamado por le tutore, o cuando hubieran pasado 15 minutos, el investigador ingresaba a la sala nuevamente y confirmaba con le tutore que hubiera concluido con su explicación. Si no, se le concedía tiempo adicional (hasta 20 minutos de tiempo total).

En la segunda parte, la interfaz de usuarie era reiniciada en modo *práctica*. Le tutore recibía el libro de ejemplos y el catálogo de imágenes y sonidos, y se le recordaba que había dos ejemplos disponibles. Luego, el robot era descubierto (si no había sido descubierto aún por le tutore) y el investigador abandonaba la sala una vez más. Cuando le tutore llamaba al investigador o cuando 40 minutos del inicio de la sesión hubieran transcurrido, la sesión se daba por finalizada.

Etapa de Revisión autónoma

Por otra parte, les niñes en el grupo *control* debían revisar el material de instrucción a solas con el pretexto de jugar aún mejor la próxima vez (en promedio 7 días después de la etapa de *Evaluación*). El procedimiento era tal que ambas etapas intermedias (de *Revisión autónoma* y de *Enseñanza entre pares*) fueran lo más similares posible, con las únicas diferencias relevantes siendo (i) social: individual vs en pares, y (ii) motivacional: revisar para mejorar el desempeño vs enseñar para permitir a une compañere jugar también.

Etapa de Evaluación final

En esta etapa, todes les niñes fueron evaluades. Corresponde a la etapa de *Evaluación* para les niñes en la condición *par-instructor*, y a la etapa de *Reevaluación* para les niñes en la condición *adulto-instructor:*

Evaluación en la condición par-instructor

Les niñes enseñades por une par fueron evaluades por primera vez, con la misma evaluación que recibieron sus compañeres en la etapa de *Evaluación inicial*. En promedio, esta etapa ocurrió 5 días después de haber sido enseñades por une par (mín=2, máx=7).

Reevaluación en la condición adulto-instructor

Por otra parte, les controles y tutores fueron evaluades por segunda vez usando la interfaz de usuarie en modo *reevaluación*. Los videos en este modo eran diferentes pero isomórficos con los videos en el modo evaluación (ver sección “3.2.1.4 Videos”). En promedio, esto ocurrió 9 días después de la *Enseñanza entre pares* para les tutores, y 8 días después de la *Revisión autónoma* para les controles.

Etapa de Seguimiento

Para evaluar el efecto de las distintas condiciones a largo plazo, realizamos un estudio complementario en el que todes les participantes fueron evaluades nuevamente dos años más tarde. Sólo se les explicó nuevamente el objetivo de la tarea y cómo usar el reproductor de videos y el navegador de ejercicios, pero no se les dio instrucciones adicionales acerca del lenguaje de programación. Además, antes de comenzar se les preguntó si recordaban el juego.

3.2.1.6 Desempeño

Para evaluar el desempeño de les participantes en las etapas de *Evaluación* y de *Reevaluación*, se consideraron dos variables relacionadas: la precisión y los tiempos de respuesta.

Precisión

La precisión fue calculada para cada ejercicio a partir de la información extraída automáticamente de los registros de video de las sesiones (ver sección “3.2.1.8 Análisis de datos”), cuantificando las diferencias entre (i) la secuencia de bloques construida por le participante, y (ii) la secuencia correcta para dicho ejercicio; es decir, la secuencia más corta que recrea la serie de eventos representada en el video (ver “Anexo 3.4: Eventos en videos”).

Para realizar esta cuantificación, utilizamos la distancia de Levenshtein (ver “Anexo 3.7: Distancias de Levenshtein”, para una descripción detallada de la estrategia empleada), que cuenta el número de pasos necesarios para transformar una secuencia en la otra; es decir, la cantidad de bloques que deben ser removidos o insertados (Levenshtein, 1966; Soukoreff & MacKenzie, 2001). En general, la secuencia al final del último intento de cada ejercicio fue tomada como la respuesta de le participante para dicho ejercicio. En caso que hubiera más de una secuencia

en el lienzo del editor, se seleccionaba aquella con la menor distancia a la secuencia correcta.

Para medir la distancia de Levenshtein entre secuencias, cada bloque de comando y de argumento fue traducido a uno de los 95 caracteres imprimibles del *Código Estándar estadounidense para el Intercambio de Información (ASCII)*, según una tabla de traducción (ver "Anexo 3.8: Tabla de traducción distancias de Levenshtein"). La distancia de Levenshtein puede tomar valores entre 0 y la suma de las longitudes de las secuencias comparadas. Para trabajar, en cambio, con valores que fueran independientes de la longitud de las secuencias, usamos el siguiente cociente *r*:

(3.1) $$r = \frac{1 - d}{long(sec1) + long(sec2)}$$,

donde *d* es la distancia de Levenshtein entre las secuencias *sec1* y *sec2*, y *long(sec)* es la longitud de la secuencia *sec*. Así, el cociente *r* varía entre 0, si las secuencias *sec1* y *sec2* no tienen ningún bloque en común, y 1, si ambas coinciden perfectamente.

Dos medidas de precisión fueron definidas: el puntaje agregado y la precisión media. El puntaje agregado es la suma de los índices de precisión de todos los ejercicios, independientemente de si fueron finalizados o no en los 20 minutos disponibles por sesión. Un ejercicio era señalado como "finalizado" si se cumplían ambas condiciones a continuación: (i) al menos un bloque fue arrastrado hacia el lienzo del editor en al menos uno de los intentos del ejercicio, y (ii) el botón "siguiente" del navegador de ejercicios fue presionado al menos una vez en dicho ejercicio. Por otro lado, la precisión media es el promedio entre los índices de precisión de los ejercicios finalizados. Como los ejercicios son de distinta dificultad, fue necesario calcular la precisión media a partir de un conjunto predefinido de ejercicios finalizado por la mayoría de les participantes. Por lo tanto, sólo los ejercicios en el bloque de ejercicios 1 a 3 fueron utilizados en el cálculo, ya que la proporción media de ejercicios finalizados en este bloque fue superior al 90% para todos los grupos experimentales (ver apartado *Proporción de ejercicios finalizados* de la sección "3.2.2.1 Efecto del tipo de instrucción").

Tiempo de respuesta

El tiempo de respuesta de un ejercicio es el tiempo invertido en todos los intentos de dicho ejercicio. Como es habitualmente el caso con los tiempos de respuesta, su distribución se encuentra sesgada (Ratcliff & McKoon, 2008), y por lo tanto se aplicó una transformación logarítmica. Para calcular el tiempo de respuesta medio, sólo los ejercicios en el bloque 1-3 fueron considerados, así como se hizo para calcular la precisión media (ver sección precedente), y las sesiones con uno o más ejercicios sin finalizar en el bloque 1-3 fueron excluidas (sesiones correspondientes a 2 tutores y a 2 aprendices).

3.2.1.7 Configuración experimental

Todas las sesiones se realizaron con una computadora portátil *Lenovo B50* con *Microsoft Windows 7*. Aunque preferíamos usar Linux por su versatilidad y capacidad de configuración para contextos experimentales, usamos Windows porque la plataforma de programación Lego WeDo no era compatible. El panel táctil de la computadora fue desactivado, y se usó un mouse óptico con cable en su lugar (pues se estimó que resultaría más sencillo de usar para les participantes). Empleamos la versión 1.2.3 de la plataforma WeDo, con la opción `CanvasLock=True` para evitar que les participantes pudieran desplazar el lienzo del editor. Para este estudio, algunos de los sonidos de la plataforma fueron reemplazados con sonidos personalizados, a fin de facilitar su representación visual. También usamos AutoHotKey[17] para restringir aún más la interacción con el sistema operativo.

Una cámara web *Logitech c920* frontal se usó para registrar las interacciones entre les participantes y con los materiales de instrucción y el catálogo de imágenes y sonidos. Otra cámara del mismo modelo, trasera, se usó para registrar la interacción de les participantes con la pantalla y el teclado de la computadora, y con el robot. La pantalla fue capturada usando el software de código abierto screen-capture-recorder.[18] El sonido ambiente se capturó usando los micrófonos integrados en la computadora y en ambas cámaras. Todos los flujos de video y audio fueron procesados y almacenados usando FFmpeg,[19] y además transmitidos mediante un

17 https://www.autohotkey.com/

18 https://github.com/rdp/screen-capture-recorder-to-video-windows-free

19 https://ffmpeg.org/

canal inalámbrico seguro al investigador ubicado fuera de la sala, para asegurar la seguridad y el bienestar de les participantes. Las acciones del reproductor de videos y del navegador de ejercicios de la interfaz de usuarie (ver sección “3.2.1.3 Interfaz de usuarie”) fueron recogidas en un registro de la sesión. Además, el detalle de la interacción con mouse y teclado fue registrado usando el software Basic Key Logger.[20]

A los fines de este estudio, se construyó una tabla de madera rectangular, con dos torres para sujetar las cámaras. En la tabla fueron marcadas las posiciones de la computadora, el mouse, el robot, los materiales impresos y las cámaras, para asegurar que fueran ubicados en la misma posición en todas las sesiones, independientemente de las posibles variaciones en la sala asignada por la escuela. El mouse y el robot fueron colocados a la derecha de la computadora. En las etapas de *Instrucción* y de *Enseñanza entre pares/Revisión*, las sesiones comenzaban con el robot escondido debajo de una caja de cartón negra. Los materiales de instrucción y el catálogo de imágenes y sonidos (con el lado correspondiente a las imágenes hacia arriba), según correspondiera, se colocaban sobre la tabla, en el espacio entre la computadora y le/s participante/s.

3.2.1.8 Análisis de datos

Para evaluar la precisión (ver apartado *Precisión* de la sección “3.2.1.6 Desempeño”) era necesario conocer la naturaleza y ubicación de los bloques en el lienzo del editor a lo largo de las sesiones en las etapas de *Evaluación* y de *Reevaluación*. Esto representaba cierta complejidad, debido al carácter cerrado de la plataforma de programación. Se consideraron diferentes alternativas para obtener esta información a partir de los registros disponibles. En primer lugar, intentamos inferirla de los registros de interacción con mouse y teclado. Sin embargo, esta estrategia resultó inapropiada, pues la respuesta de la plataforma de programación WeDo no siempre podía deducirse a partir de los datos de interacción. En segundo lugar, utilizamos estos registros para reproducir automáticamente la interacción de les participantes con la interfaz de usuarie, grabando el estado del lienzo varias veces a lo largo de cada reproducción, utilizando el formato propietario de la plataforma de programación, que logramos decodificar con técnicas de ingeniería reversa.

20 https://sites.google.com/site/basiclabbook/keyloggerbasiclabbook

Repetimos este procedimiento múltiples veces (obteniendo más de una serie de capturas por sesión) para minimizar el número de errores. Sin embargo, en algunos casos el estado final del lienzo no coincidía con el de la sesión real. Por último, la alternativa definitiva empleó los registros de pantalla: seleccionamos una captura por segundo (tomando todos los intentos de un mismo ejercicio como un único intento ininterrumpido), y luego identificamos los bloques presentes y su ubicación usando la librería para visión artificial OpenCV.[21]

El análisis estadístico de los datos y las visualizaciones se llevaron a cabo con Pandas,[22] Seaborn,[23] StatsModels,[24] SciPy,[25] afex,[26] y otras librerías de Python y R, en Jupyter Notebook.[27] Se verificó el cumplimiento de los supuestos para las pruebas estadísticas paramétricas utilizadas, y se usaron pruebas no paramétricas donde estos supuestos no se cumplían.

3.2.2 Resultados

3.2.2.1 Efecto del tipo de instrucción

En este estudio, algunes de les participantes fueron enseñades por une instructore adulte en una clase grupal (condición *adulto-instructor*), mientras que otres fueron enseñades por une par en una clase une-a-une (condición *par-instructor*). En esta sección evaluamos los efectos diferenciales de estos dos tipo de instrucción, sobre (1) el desempeño, (2) las estrategias de resolución, y (3) la confianza.

Tipo de instrucción y desempeño

En primer lugar, evaluamos los efectos sobre el desempeño. Concretamente, evaluamos (1) el puntaje agregado, (2) la proporción de ejercicios finalizados, (3) el tiempo de respuesta por ejercicio, y (4) la precisión por ejercicio. Además, consideramos una condición control adicional *sin-instrucción* (5).

21 https://opencv.org

22 https://pandas.pydata.org/

23 https://seaborn.pydata.org/

24 https://www.statsmodels.org/

25 https://www.scipy.org

26 https://afex.singmann.science/

27 https://jupyter.org/)

Puntaje agregado

Primero comparamos el puntaje agregado en la etapa de *Evaluación* entre condiciones; es decir, la suma de los índices de precisión de todos los ejercicios, independientemente de si fueron finalizados o no, en los 20 minutos disponibles (ver apartado *Precisión* de la sección “3.2.1.6 Desempeño”). Realizamos un análisis de la covarianza (ANCOVA) de dos vías, con condición (*adulto-instructor* vs *par-instructor*) y género como factores, y edad como covariable. Les niñes que fueron enseñades por une par mostraron puntajes similares, IC 68% [3,80, 4,61], a los de les niñes enseñades por une adulte, IC 68% [3,62, 4,09], $F(37, 1) = 0{,}510$, $p = 0{,}480$ (Figura 3.3).

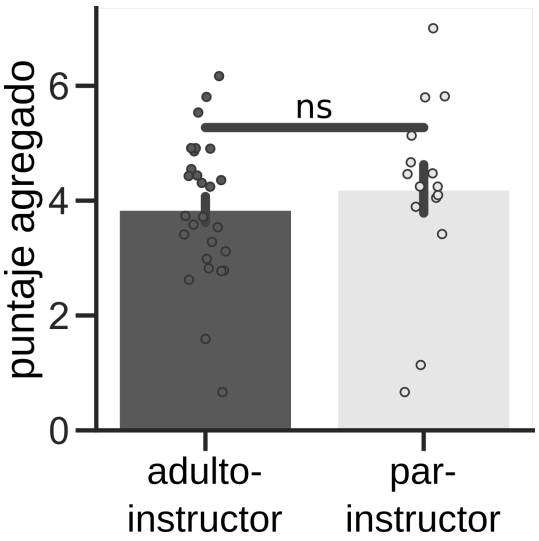


Figura 3.3: *Puntaje agregado*

El puntaje agregado es la suma de los índices de precisión los 10 ejercicios, incluyendo los no intentados (con índice de precisión = 0). Les niñes que aprendieron de une par (condición *par-instructor*) mostraron un desempeño similar al de les niñes que aprendieron de une adulte (condición *adulto-instructor*). ns: no significativo. Las barras de error representan intervalos de confianza *bootstrap* del 68%.

Por otra parte, encontramos un efecto principal de *género* sobre el puntaje agregado, con los puntajes de las niñas en promedio menores, IC 68% [3,39, 3,90], que el de los niños, IC 68% [4,20, 4,85], $F(37, 1) = 5{,}06$, $p = 0{,}031$. Sin embargo, no observamos efectos significativos de la interacción *condición:género*, $F(37, 1) = 0{,}256$, $p = 0{,}616$.

Proporción de ejercicios finalizados

Luego, se procedió a analizar el desempeño ejercicio por ejercicio. En los 20 minutos disponibles en la etapa de *Evaluación*, les niñes podían pasar tanto tiempo como quisieran en cualquiera de los 10 ejercicios, afectando la proporción de ejercicios *finalizados*. Además, los ejercicios diferían en complejidad (con los últimos ejercicios siendo más complejos que los primeros) y es de suponer que esta complejidad haya influido de manera diferente en los índices de precisión. Por lo tanto, para poder comparar la precisión media entre sujetos correctamente, sin sesgos, es importante usar solamente los ejercicios que hubieran sido finalizados por la mayoría de les participantes en todas las condiciones. Por este motivo, los ejercicios fueron divididos en tres bloques: 1-3, 4-7 y 8-10.

En promedio, les participantes finalizaron casi todos los ejercicios en el bloque 1-3 (en promedio, 94% de ejercicios finalizados en la condición *adulto-instructor*, y 98% en la condición *par-instructor*), y por lo tanto este bloque fue seleccionado para estudiar el desempeño ejercicio por ejercicio. En el bloque 4-7, en cambio, no sólo la proporción media de ejercicios finalizados fue menor, sino que además resultó significativamente más baja en la condición *adulto-instructor*, IC 68% [0,33, 0,45], que en la condición *par-instructor*, IC 68% [0,57, 0,73], como muestra una prueba Kruskal-Wallis, $H(1) = 5{,}51$, $p = 0{,}019$ (Figura 3.4A, panel izquierdo). Además, esta proporción resultó significativamente menor para las niñas, IC 68% [0,33, 0,44], que para los niños, IC 68% [0,58, 0,73], como muestra otra prueba Kruskal-Wallis, $H(1) = 6{,}32$, $p = 0{,}012$.

Tiempo de respuesta por ejercicio

Consistentemente, les niñes en la condición *par-instructor* pasaron significativamente menos tiempo por ejercicio (segundos) en el bloque 1-3, n=13, IC 68% [114, 138], que les niñes en la condición *adulto-instructor*, n=25, IC 68% [172, 201], como muestra un ANCOVA de tres vías con medidas repetidas, con condición y género como factores inter-sujeto, ejercicio como factor intra-sujeto, y edad como covariable, $F(1, 33) = 7{,}60$, $p = 0{,}009$ (Figura 3.4A, panel derecho). El hecho de que les niñes enseñades por une par pasaron menos tiempo por ejercicio (alrededor de 60 segundos menos) en el bloque 1-3 podría explicar por qué finalizaron más ejercicios en el bloque 4-7.

Además, el análisis descriptivo muestra una posible asociación entre género y el tiempo de respuesta medio, con las niñas dedicando más tiempo (en segundos) por ejercicio, n=23, IC 68% [166, 193], que los niños, n=15, IC 68% [125, 156]; pudiendo explicar en parte que éstos finalizaran más ejercicios en el bloque 4-7 que aquellas. Sin embargo, este efecto no es significativo, $F(33, 1) = 3{,}14$, $p = 0{,}085$.

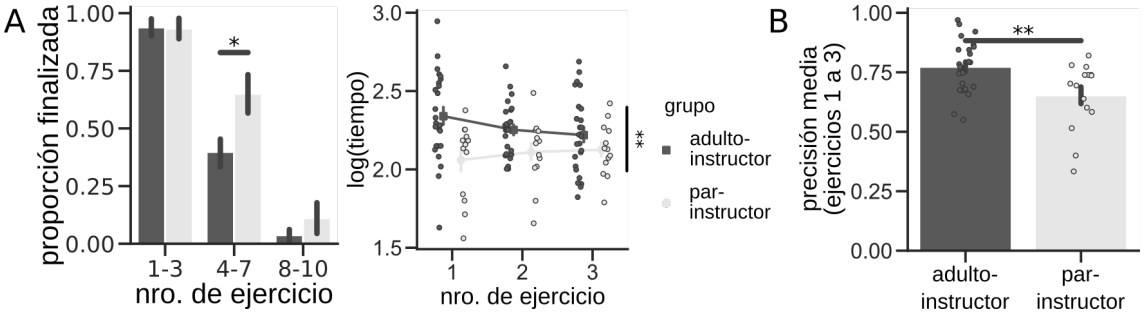


Figura 3.4: *Desempeño por ejercicio*

(A) Izquierda. Proporción de ejercicios finalizados. Les niñes que aprendieron de une par finalizaron más ejercicios en la etapa de evaluación que les niñes que aprendieron de une adulte.

Derecha: Tiempo por ejercicio. Les niñes que aprendieron de une par dedicaron menos tiempo por ejercicio que les niñes que aprendieron de une adulte. Los tiempos de respuesta siguen una distribución log-normal, y por lo tanto se les aplicó una transformación logarítmica. Les participantes que no finalizaron todos los ejercicios 1 a 3 fueron excluides.

(B) Precisión por ejercicio. Les niñes que aprendieron de une par fueron menos precisos que les que aprendieron de une adulte. Une participante que no finalizó ninguno de los ejercicios 1 a 3 fue excluide.

(*: $p < 0{,}05$, **: $p < 0{,}01$. Las barras de error representan intervalos de confianza *bootstrap* del 68%).

Precisión por ejercicio

Sin embargo, les niñes enseñades por sus pares mostraron significativamente menor precisión por ejercicio, IC 68% [0,62, 0,69], que les niñes enseñades por une adulte, IC 68% [0,75, 0,79], en el bloque 1-3, como indica un ANCOVA de dos vías, con condición y género como factores, y edad como covariable, $F(36, 1) = 9{,}07$, $p = 0{,}005$. Esta compensación entre velocidad y precisión, con tiempos de respuesta

más cortos -que resultan en mayor proporción de ejercicios finalizados (Figura 3.4A)- contrarrestados por menores índices de precisión (Figura 3.4B), podría explicar por qué el puntaje agregado de les niñes enseñades por une par resultó indistinguible del de les niñes enseñades por une adulte (Figura 3.3).

Por otra parte, el análisis no encontró efecto principal de género sobre la precisión media, que no resultó significativamente diferente entre niñas, IC 68% [0,69, 0,75], y niños, IC 68% [0,72, 0,77], $F(36, 1) = 0{,}510$, $p = 0{,}480$.

Condición sin-instructor

En las secciones precedentes identificamos similitudes y diferencias entre aprender de une adulte y aprender de une par. Sin embargo, no podemos descartar que les participantes podrían haber aprendido también de la propia etapa de *Evaluación*, por prueba y error. En extremo, es posible que intentar aprender de une par fuera en vano, y que esta condición haya sido equivalente a aprender jugando, sin instrucción previa.

Por lo tanto, realizamos un estudio complementario, en el que se introdujo una tercera condición en la que les participantes fueron evaluades sin haber sido enseñades por une adulte ni por une par. El estudio se llevó a cabo durante un evento de ciencia abierta organizado en el Centro Cultural de la Ciencia (C3) en la ciudad de Buenos Aires. En el evento, les niñes llegaban con sus familias, y podían participar de distintas actividades hasta que decidieran retirarse. A diferencia del estudio principal, que se realizó en dos segundos grados de una escuela primaria, con participantes de entre 7 y 8 años de edad, la edad de les participantes en este estudio varió más ampliamente. El número de participantes entre 7 y 8 años resultó insuficiente, y por lo tanto incluimos los datos de les participantes entre 7 y 10 años de edad.

El estudio incluyó dos condiciones experimentales. En una, une adulte instructore explicó a les participantes cómo jugar, como en el estudio principal (condición *adulto-instructor*); en la otra, se les pidió a les participantes que jugaran sin instrucción previa (condición *sin-instructor*). En total participaron 11 niñes enseñades por une adulte y 9 niñes sin instrucción. Además, se incluyeron los datos de les 27 niñes enseñades por une adulte y de les 15 enseñades por une par del estudio principal. Si bien la decisión de comparar datos provenientes de estudios diferentes

no es óptima (lo ideal sería haber realizado un nuevo estudio con las tres condiciones), el hecho de que los datos de la condición *adulto-instructor* (la única repetida en ambos estudios) resultaran indistinguibles entre estudios como veremos mas adelante, ayuda a justificar esta decisión.

En esta versión del estudio, la etapa de *Evaluación* se realizó a continuación de la etapa de *Instrucción* (si correspondía) y tuvo que ser modificada por limitaciones materiales y del evento. Puntualmente, les participantes tenían 10 minutos en lugar de 20, no se incluyeron ejercicios que implicaran rotar el motor, y el total de ejercicios fue 6 en vez de 10 (aunque los ejercicios 1 a 4 eran idénticos entre ambos estudios). Por lo tanto, para comparar la precisión media entre condiciones, sólo los ejercicios 1 y 2 fueron considerados porque sólo éstos fueron finalizados por la mayoría de les participantes; concretamente, por al menos dos terceras partes de les participantes en cada condición. Los datos de les niñes que no finalizaron ambos ejercicios 1 y 2 fueron excluidos. Al final, los datos de 36 niñes en la condición *adulto-instructor* (9 de este estudio y 25 del estudio principal), 14 en la condición *par-instructor* (del estudio principal), y 6 en la condición *sin-instructor* (de este estudio) fueron incluidos en el análisis.

La precisión media en el bloque de ejercicios 1-2 de les niñes en la condición *adulto-instructor* (la única condición repetida entre ambos estudios) no resultó significativamente diferente entre este estudio complementario, IC 68% [0,69, 0,79], y el estudio principal, IC 68% [0,77, 0,81], como se observa con una prueba *t* para dos muestras independientes, $p = 0{,}352$ (cabe aclarar que la media del estudio principal reportada aquí no coincide exactamente con la reportada en el apartado *Precisión por ejercicio* de esta sección, porque allí se consideró el bloque de ejercicios 1-3, y aquí el bloque 1-2). Por lo tanto, los datos de les niñes en la condición *adulto-instructor* de ambos estudios se juntaron para los análisis subsiguientes.

Para comparar la precisión media entre condiciones, un ANCOVA de dos vías, con condición y género como factores, y edad como covariable, mostró un efecto principal de condición en la precisión media, $F(51, 2) = 12{,}6$, $p = 3{,}69\text{E-}05$. Una prueba post-hoc HSD de Tukey mostró que les participantes sin instrucción de este estudio complementario fueron significativamente menos precisos, IC 68% [0,44, 0,56], que les participantes enseñades por une par del estudio principal, IC 68%

[0,60, 0,67], p = 0,045; sugiriendo que ser enseñade por une par no es lo mismo que no ser enseñade en absoluto. Por otra parte, como era esperable del estudio principal, la precisión media de les niñes en la condición *par-instructor*, IC 68% [0,60, 0,67], fue significativamente más baja en este bloque 1-2 que la de les niñes en la condición *adulto-instructor*, IC 68% [0,75, 0,80], como indica una prueba post-hoc HSD de Tukey, p = 0,007.

Tipo de instrucción y estrategias de resolución

Otro aspecto que exploramos en función de quién fue le instructore es el de los distintos comportamientos y estrategias exhibidos, independientemente del desempeño. En particular nos enfocamos en un comando que les niñes debían incluir en sus programas para borrar imágenes y textos en la pantalla virtual del entorno programable al final de todos los ejercicios, para reiniciarla. Para lograr esto, les niñes debían usar el bloque de comando correspondiente (de texto o de imagen) y acoplarle un bloque de argumento vacío (ver “Anexo 3.4: Eventos en videos”). Este comportamiento fue demostrado por el instructor en la etapa de *Instrucción* de la condición *adulto-instructor*.

Notablemente, mientras la mitad (52%) de les niñes enseñades por une adulte intentaron reiniciar la pantalla virtual en al menos uno de los ejercicios que así lo requerían, ningune de les niñes enseñades por une par mostraron este comportamiento. Para saber si esta diferencia de comportamiento influyó en las diferencias de desempeño observadas, realizamos un análisis complementario ignorando este tipo de combinaciones de bloques. Sin embargo, excluirlas no fue suficiente para anular la diferencia de precisión por ejercicio entre condiciones descripta en la sección precedente.

Tipo de instrucción y confianza

Como introdujimos en el Capítulo 1, además del posible efecto sobre el desempeño, la enseñanza entre pares podría impactar sobre habilidades cognitivas, como la metacognición: la capacidad de evaluar y reflexionar acerca del propio conocimiento. Para evaluar este aspecto, en este estudio analizamos la confianza de les participantes sobre su propio desempeño.

Al final de cada ejercicio, les participantes debían responder si habían logrado reproducir el video propuesto, pudiendo responder “sí”, “más o menos” o “no” (ver apartado *Pantalla de autoevaluación* de la sección “3.2.1.3 Interfaz de usuarie”). En este análisis nos enfocamos en la proporción de ejercicios evaluados como “sí” (alta confianza) con respecto al total de ejercicios *finalizados* en el bloque de ejercicios 1 a 3, teniendo en cuenta sesiones con al menos un ejercicio autoevaluado en este bloque (es decir, con hasta dos ejercicios sin autoevaluación).

Concretamente, comparamos la proporción de ejercicios evaluados con alta confianza entre niñes que aprendieron de une par, IC 68% [0,64, 0,84], y niñes que aprendieron de une adulte, IC 68% [0,77, 0,84]. Sin embargo, no encontramos diferencias significativas entre ambas condiciones, como muestra una prueba no paramétrica de Mann-Whitney, $U = 162$, $p = 0{,}171$.

En resumen, desde la perspectiva del puntaje agregado –es decir, la suma de los índices de precisión de todos los ejercicios– les niñes que aprendieron de une par resultaron indistinguibles de les que aprendieron de une adulte. No obstante, una mirada más cercana revela que les niñes enseñades por une par fueron más rápides, invirtiendo menos tiempo por ejercicio y finalizando más ejercicios por sesión. Esto ocurrió a expensas de una menor precisión por ejercicio, aunque reportando niveles de confianza semejantes. Notablemente, esta condición de aprendizaje facilitado por une par no sería simplemente sustituible por una condición sin instrucción.

Análogamente, aunque las niñas mostraron menor puntaje agregado que los niños, una mirada más cercana revela que esta diferencia no se debió a una menor precisión por ejercicio, sino a que fueron más lentas y por lo tanto finalizaron menos ejercicios en el tiempo disponible.

Asimismo, encontramos diferencias entre condiciones en al menos una estrategia de resolución, pero no serían responsables (al menos no únicamente) de las diferencias de desempeño observadas.

3.2.2.2 Efecto de enseñar

La segunda hipótesis proponía que les niñes se verían beneficiades de su rol como tutores. Para contrastar esta hipótesis, luego de observar a les niñes enseñades por une adulte en la etapa de *Evaluación*, a una mitad (grupo *tutor*) se les pidió que enseñaran a une compañere inexperte, mientras que a la otra mitad (grupo *control*) se les pidió que revisaran los materiales de instrucción a solas. Finalmente, ambos grupos fueron evaluados nuevamente en la etapa de *Reevaluación* (ver sección "3.2.1.5 Procedimiento"). Consideramos los efectos sobre (1) el desempeño y (2) la confianza.

Enseñanza y desempeño

Puntaje agregado

Para comparar los puntajes agregados entre etapas (*Evaluación* y *Reevaluación*) y grupos (*control* y *tutor*), se hizo un ANCOVA de tres vías con medidas repetidas, con grupo y género como factores inter-sujeto, etapa como factor intra-sujeto, y edad como covariable. Encontramos un efecto significativo de la etapa, con puntajes agregados más altos en la etapa de *Reevaluación* (controles: IC 68% [5,44, 6,04], tutores: IC 68% [4,84, 5,68]) que en la de *Evaluación* (controles: IC 68% [3,83, 4,25], tutores: IC 68% [3,33, 4,09]), $F(1, 22) = 55{,}0$, $p = 2{,}04\text{E-}07$. No obstante, no observamos un efecto significativo de interacción entre etapa y grupo, $F(1, 22) = 0{,}207$, $p = 0{,}654$ (Figura 3.5). Por lo tanto, el aumento en el puntaje agregado no parecería haber surgido como resultado de enseñar a une par, sino simplemente como un efecto genérico de reevaluación y/o motivado por la revisión de los materiales (ya sea de a pares o a solas).

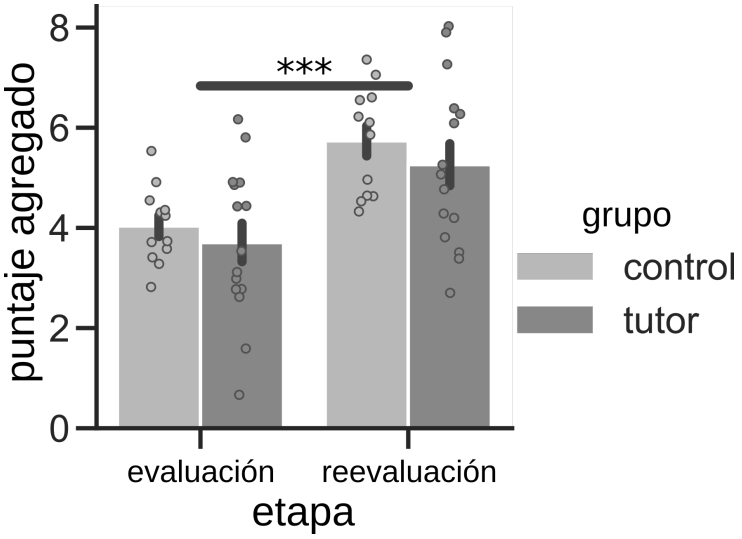


Figura 3.5: *Efecto de enseñar*

Tanto controles como tutores mejoraron en la etapa de *Reevaluación*, pero no encontramos evidencia de que enseñar haya provisto un beneficio (o perjuicio) adicional a les tutores en su desempeño. (***: p < 0,001. Las barras de error representan intervalos de confianza *bootstrap* del 68%).

Para entender mejor las razones detrás de este cambio, realizamos un análisis más fino del desempeño, de forma similar al que se hizo en la etapa de *Evaluación* para comparar mejor las condiciones *adulto-instructor* y *par-instructor* (ver sección “3.2.2.1 Efecto del tipo de instrucción”). Con este objetivo, dividimos el análisis del puntaje agregado en la etapa de *Reevaluación* en (i) la proporción de ejercicios finalizados por bloque de ejercicios (es decir, bloques 1-3, 4-7 y 8-10), (ii) el tiempo medio por ejercicio en el bloque 1 a 3, y (iii) la precisión media por ejercicio en el bloque de ejercicios 1 a 3.

Proporción de ejercicios finalizados

Un ANCOVA de dos vías con medidas repetidas, con grupo y género como factores inter-sujeto, etapa como factor intra-sujeto, y edad como covariable, encontró un efecto significativo de la etapa sobre la proporción de ejercicios finalizados, $F(22, 1) = 13{,}7$, $p = 0{,}001$, sin efecto significativo de grupo, $F(22, 1) = 1{,}80$, $p = 0{,}19$, o de la interacción grupo:etapa, $F(22, 1) = 0{,}09$, $p = 0{,}77$.

En promedio, les niñes finalizaron significativamente más ejercicios (1,18 ejercicios más) en el bloque 4-7 de la *Reevaluación* (*control*: IC 68% [0,73, 0,85], *tutor*: IC 68%

[0,55, 0,72]), que en el mismo bloque de la *Evaluación* (*control*: IC 68% [0,38, 0,54], *tutor*: IC 68% [0,27, 0,43]).

Tiempo de respuesta por ejercicio

Un ANCOVA de dos vías con medidas repetidas, con grupo y género como factores inter-sujeto, etapa como factor intra-sujeto, y edad como covariable, encontró un efecto significativo de la etapa sobre el tiempo medio por ejercicio, $F(20, 1) = 7{,}58$, $p = 0{,}01$, sin efecto significativo de grupo, $F(20, 1) = 0{,}40$, $p = 0{,}53$, o de la interacción grupo:etapa, $F(20, 1) = 0{,}17$, $p = 0{,}69$.

En promedio, les niñes fueron significativamente más rápides por ejercicio (63 segundos) en la *Reevaluación* (*control*: IC 68% [118, 137], *tutor*: IC 68% [105, 138]), que en la *Evaluación* (*control*: IC 68% [166, 207], *tutor*: IC 68% [168, 209]), en el bloque de ejercicios 1-3.

Precisión por ejercicio

Un ANCOVA de dos vías con medidas repetidas, con grupo y género como factores inter-sujeto, etapa como factor intra-sujeto, y edad como covariable, encontró un efecto significativo de la etapa sobre la precisión media por ejercicio, $F(21, 1) = 4{,}99$, $p = 0{,}04$, sin efecto significativo de grupo, $F(21, 1) = 2{,}10$, $p = 0{,}16$, o de la interacción grupo:etapa, $F(21, 1) = 0{,}01$, $p = 0{,}92$.

En promedio, la precisión media por ejercicio en el bloque 1-3 fue apenas 0,070 puntos más alta en la etapa de *Reevaluación* (*control*: IC 68% [0,81, 0,86], *tutor*: IC 68% [0,83, 0,88]), que en la de *Evaluación* (*control*: IC 68% [0,74, 0,79], *tutor*: IC 68% [0,75, 0,81]).

Enseñanza y confianza

En este análisis comparamos la diferencia entre etapas en la proporción de ejercicios respondidos con alta confianza para ambos grupos, *control* y *tutor*. Excluimos une participante (del grupo *tutor*) que no había finalizado ningún ejercicio en el bloque 1-3 en al menos una de las dos etapas (*Evaluación* o Reevaluación).

Debido a la distribución no normal de los valores de esta variable, realizamos dos pruebas no paramétricas de Wilcoxon (de los rangos con signo) independientes, una por cada grupo, con valores *p* corregidos por el método de Bonferroni para comparaciones múltiples. Los resultados muestran que, para el grupo *control* (n=12),

la proporción de ejercicios respondidos con alta confianza fue significativamente mayor en la etapa de *Reevaluación*, IC 68% [0,89, 1,00], que en la de *Evaluación*, IC 68% [0,67, 0,75], p = 0,008, acompañando el mejor desempeño observado en esta etapa. No obstante, esta proporción no cambió significativamente en el grupo *tutor* (n=14, *Reevaluación*: IC 68% [0,81, 0,90], *Evaluación*: IC 68% [0,62, 0,81], p = 0,151), a pesar del aumento equivalente en el desempeño (Figura 3.6).

Esta diferencia, en apariencia paradójica, con participantes del grupo *control* demostrando mayor confianza que les del grupo *tutor*, podría en realidad sugerir una mejor metacognición en estos últimos, como proponemos en la sección “3.2.3 Discusión”.

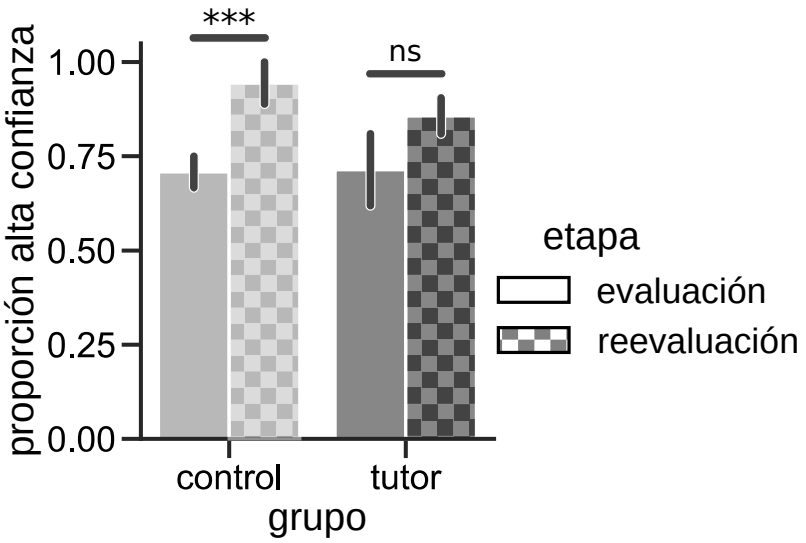


Figura 3.6: *Confianza*

Enseñar a une par podría haber modulado la confianza sobre el propio desempeño. Si bien ambes controles y tutores demostraron un mejor desempeño en la *Reevaluación* (ver Figura 3.5), sólo les controles demostraron un aumento significativo de la confianza en esta etapa. (ns: no significativo; **: p < 0,01. Las barras de error representan intervalos de confianza *bootstrap* del 68%).

En resumen, aunque les niñes mejoraron de la *Evaluación* a la *Reevaluación*, esta mejoría no estuvo relacionada con el hecho de enseñar a une par y por lo tanto fue la misma para tutores que para controles. Un análisis detallado reveló que este aumento en el puntaje agregado en ambos grupos se debió no sólo a un pequeño incremento en la precisión media, sino fundamentalmente a una disminución en los

tiempos de respuesta por ejercicio, que en consecuencia derivó en una mayor proporción de ejercicios finalizados. Enseñar a une par, no obstante, parecería haber modulado la confianza sobre el propio desempeño, sugiriendo un posible impacto sobre la capacidad metacognitiva.

3.2.2.3 Efectos a largo plazo

Para probar la posibilidad de que enseñar a une par podría tener efectos a largo plazo, realizamos un estudio complementario (ver apartado *Etapa de Seguimiento* de la sección "3.2.1.5 Procedimiento") en el que todos los grupos del estudio principal (es decir, sin incluir les participantes del estudio complementario descripto en el apartado *Condición sin-instructor* de la sección "3.2.2.1 Efecto del tipo de instrucción") fueron evaluades nuevamente dos años más tarde. De les 42 participantes originales (*control*=12, *tutor*=15, *aprendiz*=15), 35 se encontraban en condiciones para participar de esta etapa de *Seguimiento* dos años después (*control*=8, *tutor*=14, *aprendiz*=13). Nuevamente, consideramos los efectos sobre (1) el desempeño y (2) la confianza.

Desempeño a largo plazo

Primero, la mayoría de les niñes (72%) respondieron que recordaban el juego, sin diferencias significativas entre grupos, como se muestra con una prueba Chi cuadrado de independencia de variables, p = 0,282.

Luego, para analizar los puntajes agregados, realizamos un ANCOVA de tres vías con medidas repetidas, con grupo (*control*, *tutor* y *aprendiz*) y género como factores inter-sujeto, etapa (*Evaluación* y *Seguimiento*) como factor intra-sujeto, y edad como covariable. Encontramos un efecto principal de *etapa* sobre el puntaje agregado, con puntajes en la etapa de *Seguimiento*, IC 68% [1,86, 2,33], significativamente menores que en la etapa de *Evaluación*, IC 68% [3,73, 4,22], $F(28, 1) = 31{,}9$, p = 4,79E-06. Por otra parte, no observamos efecto principal de *grupo*, $F(28, 2) = 1{,}01$, p = 0,377, o de la interacción *grupo:etapa*, $F(28, 2) = 0{,}913$, p = 0,413, con puntajes agregados en la etapa de *Seguimiento* indistinguibles entre los grupos *control*, IC 68% [2,06, 3,26], *tutor*, IC 68% [1,56, 2,31], y *aprendiz*, IC 68% [1,64, 2,20]. La etapa de *Reevaluación* no fue incluida en el análisis porque el grupo *aprendiz* no participó de ella.

A continuación, para un análisis más fino del desempeño, consideramos la precisión en el bloque de ejercicios 1-3. Sólo incluimos los datos de les participantes que finalizaron al menos un ejercicio en el bloque 1-3 en ambas etapas, *Evaluación* y *Seguimiento* (*control*=7, *tutor*=12, *aprendiz*=12). Realizamos un ANCOVA de tres vías con medidas repetidas, con grupo y género como factores inter-sujeto, etapa como factor intra-sujeto, y edad como covariable. Encontramos un efecto principal de *etapa*, con precisiones medias en la etapa de *Seguimiento*, IC 68% [0,31, 0,37], significativamente menores que en la etapa de *Evaluación*, IC 68% [0,68, 0,75], $F(24, 1) = 76,3$, $p = 6,47E-09$. Además, un ANCOVA de dos vías en la etapa de *Seguimiento* solamente, con grupo y género como factores, y edad como covariable, no encontró efecto principal de *grupo* en la precisión media, $F(24, 2) = 2,26$, $p = 0,126$, sin diferencias significativas en esta etapa entre grupos *control*, IC 68% [0,39, 0,52], *tutor*, IC 68% [0,26, 0,35], y aprendiz, IC 68% [0,26, 0,34].

Finalmente, comparamos la precisión media en la etapa de *Seguimiento* con la de niñes evaluades directamente, sin instrucción previa (condición *sin-instructor*). En este caso, sólo se consideraron los ejercicios 1 y 2 (ver apartado *Condición sin-instructor* de la sección “3.2.2.1 Efecto del tipo de instrucción”). Únicamente los datos de les niñes que finalizaron ambos ejercicios 1 y 2 fueron incluidos (*control*=7, *tutor*=11, *aprendiz*=11, *sin-instructor*=6). Realizamos un ANCOVA de dos vías, con grupo y género como factores, y edad como covariable, pero no se observó efecto principal de *grupo* en la precisión media, $F(26, 3) = 0,742$, $p = 0,537$, sin diferencias significativas entre los grupos *control*, IC 68% [0,32, 0,52], *tutor*, IC 68% [0,27, 0,37], *aprendiz*, IC 68% [0,23, 0,33], y condición *sin-instructor*, IC 68% [0,42, 0,56].

Confianza a largo plazo

Por otra parte, para evaluar un posible efecto de grupo sobre la confianza a largo plazo, analizamos la diferencia en la proporción de ejercicios respondidos con alta confianza (en el bloque de ejercicios 1 a 3) entre esta etapa de *Seguimiento*, y la última etapa del estudio principal; es decir, la etapa de *Evaluación* para el grupo *aprendiz*, y la etapa de *Reevaluación* para los grupos *control* y *tutor*. Nos enfocamos en el bloque de ejercicios 1 a 3, y se consideraron los datos de les participantes que reportaron confianza en al menos un ejercicio de este bloque en ambas etapas (*Evaluación/*Reevaluación y *Seguimiento*).

Con este fin, realizamos tres pruebas no paramétricas de Wilcoxon (de los rangos con signo) independientes, una para cada grupo, con corrección de valores *p* por método Bonferroni para comparaciones múltiples, debido a la distribución no normal de la variable estudiada. De forma similar a lo observado entre las etapas de *Evaluación* y *Reevaluación* (ver apartado *Enseñanza y confianza* de la sección "3.2.2.2 Efecto de enseñar"), mientras el grupo *tutor* (n=12) redujo significativamente sus niveles de confianza en esta etapa de *Seguimiento*, IC 68% [0,39, 0,57], con respecto a la etapa de *Reevaluación*, IC 68% [0,83, 0,94], p = 0,035, acompañando la disminución de desempeño observada, esta reducción no resultó significativa en el grupo *control* (n=7, *Seguimiento*: IC 68% [0,81, 1,00], *Reevaluación*: IC 68% [0,57. 0,83], p = 0,066), a pesar de una disminución de desempeño similar. El grupo *aprendiz* (n=12) también demostró una reducción significativa de la confianza en la etapa de *Seguimiento*, IC 68% [0,22, 0,39], con respecto a la etapa de *Evaluación*, IC 68% [0,53, 0,78], p = 0,045.

En resumen, si bien la mayoría de les participantes declararon recordar el juego en la etapa de *Seguimiento*, dos años después, el desempeño fue el mismo para todos los grupos, menor que en la etapa de *Evaluación*, e indistinguible de niñes que no fueron enseñades en absoluto. No obstante, sólo les individues que participaron de la *Enseñanza entre pares* (grupos *tutor* y *control*) parecerían dar cuenta de este menor desempeño, reduciendo sus niveles de confianza, a diferencia de quienes participaron de la *Revisión autónoma* (grupo *control*).

3.2.3 Discusión

En este estudio abordamos la problemática de enseñar programación a niñes considerando el posible rol de la enseñanza entre pares. Investigamos tanto si un grupo de niñes podría aprender a programar de une par, como así también si éstes se beneficiarían por enseñar a otres. Para hacerlo, analizamos su desempeño en una serie de tareas de programación.

3.2.3.1 Aprendiendo de une par

Para evaluar si es posible que une niñe aprenda a programar de une par, primero comparamos el puntaje agregado entre niñes enseñades por une adulte y niñes

enseñades por une par. Este tipo de evaluaciones son habitualmente utilizadas en contextos escolares, donde les niñes reciben un cierto número de preguntas y tienen un tiempo fijo para responderlas. Al final, los puntajes de las preguntas individuales se suman, de modo que no importa solamente cuán bien respondieron cada pregunta, sino también cuántas lograron responder en el tiempo disponible. En este escenario, una pregunta no respondida por falta de tiempo recibe cero puntos, del mismo modo que una pregunta no respondida por falta del conocimiento necesario. De forma análoga, la primera estrategia que adoptamos para medir desempeño comprendió sumar los índices de precisión de 10 ejercicios de una tarea de programación, sin importar si les niñes habían conseguido finalizarlos (siquiera intentarlos) o no en los 20 minutos que disponían. Notablemente, según esta primera medida general de desempeño, les niñes que fueron enseñades por une par se desempeñaron tan bien como aquelles enseñades por une adulte.

Pero este resultado no significa necesariamente que les niñes enseñades por une adulte y les enseñades por une par sean indistinguibles. Por consiguiente, a continuación usamos un enfoque más fino para medir desempeño, examinando precisión y tiempo de respuesta por separado. Con este enfoque, observamos diferencias en el equilibrio entre velocidad y precisión: les niñes enseñades por une par parecen haber adoptado una estrategia menos conservadora, con menores tiempos de respuesta que sus compañeres enseñades por une adulte, pero a expensas de una menor precisión. Cabe destacar que la precisión de les niñes en la condición *adulto-instructor* podría haber sido mayor si sus clases hubieran sido individuales (Bloom, 1984), como las de la condición *par-instructor*, ampliando aún más la brecha.

Este hallazgo es curiosamente similar al de un estudio previo que muestra que jugadores de ajedrez, cuando se enfrentan a contrincantes más fuertes, adoptan una estrategia de juego más conservadora (Slezak & Sigman, 2012), y puede interpretarse en el marco de la teoría del enfoque regulatorio (Higgings, 1997). Según esta teoría, las personas se pueden situar en un modo de *promoción*, en el que el enfoque está en alcanzar un estado deseado (como ganar un juego de ajedrez o, en este caso, resolver exitosamente una tarea de programación), o en un modo de *prevención*, en el que el foco está en evitar estados no deseados (por ejemplo, cometer errores). En el modo de *promoción* las personas, ansiosas por

alcanzar el estado deseado, aumentan su velocidad disminuyendo su precisión (Slezak & Sigman, 2012).

En el marco de esta teoría, les niñes enseñades por une adulte, que aprendieron de una figura de autoridad con más conocimiento del tema, parecen haber adoptado el modo de *prevención*, reflejado en una forma más cuidadosa de jugar, evitando errores. Mientras que les niñes enseñades por une par podrían haber tomado, en cambio, el modo de *promoción*, asumiendo más riesgos por prueba y error. Si bien las decisiones apresuradas a menudo derivan en elecciones subóptimas, las decisiones más precisas no son efectivas si demoran demasiado tiempo (Standage et al., 2015). Por ejemplo, adoptar una estrategia conservadora frente a un adversario más fuerte en ajedrez en realidad reduce las probabilidades de salir victorioso (Slezak & Sigman, 2012). Entonces, ¿cuál es el mejor balance entre velocidad y precisión? Probablemente dependa de la situación entre manos si preferir métodos de enseñanza que promuevan la opción conservadora del modo de *prevención*, con un enfoque en evitar errores, o métodos que promuevan la opción menos conservadora del modo de *prevención*, que se enfoca en los resultados y admite cierta prueba y error.

En resumen, así cómo la instrucción puede cambiar las políticas exploratorias en niñes, limitando la exploración espontánea y el descubrimiento (Bonawitz et al., 2011), nuestros resultados podrían sugerir una suerte de interacción entre la enseñanza entre pares y las creencias epistemológicas de les niñes, que podrían haber desarrollado nociones diferentes acerca de cómo utilizar los nuevos conocimientos según quién les enseñó. Es posible que esto haya sido el producto de distintas expectativas acerca de para qué estaban aprendiendo, primero, y para qué estaban jugando, después. Tal vez les niñes enseñades por une adulte suponían estar siendo evaluades, mientras que les enseñades por une par suponían que estaban jugando simplemente. No obstante, hacen falta estudios adicionales, que manipulen las expectativas de les participantes, para poder evaluar esta posibilidad.

Por otra parte, también observamos diferencias en las estrategias de resolución. El hecho que la mitad de les niñes enseñades por une adulte hayan empleado una estrategia demostrada por su instructor, mientras que ningune de les enseñades por une par la hayan exhibido, sugiere que esta estrategia podría no haber sido enseñada por les niñes instructores, compatible con la idea de que éstes no se

limitan a imitar a su instructore adulte, sino que por el contrario procesan la información que se les ha suministrado y elaboran una forma de enseñar propia (Bridgers et al., 2019). No obstante, será necesario conducir estudios adicionales de las interacciones registradas para contrastar esta hipótesis.

3.2.3.2 El efecto de enseñar a une par

Para poder evaluar el posible efecto de enseñar, nos aseguramos que las etapas de *Enseñanza entre pares* y de *Revisión autónoma* fueran tan comparables como fuese posible. Así, las únicas diferencias destacables entre ambas fueron (i) social: en un caso les niñes interactuaban con une par, mientras que en el otro estaban a solas; y (ii) motivacional: en un caso el objetivo era enseñar a une compañere para que éste pudiera jugar también, mientras que en el otro era revisar el material para jugar aún mejor la próxima vez. Nuestros resultados muestran que les niñes mejoraron su puntaje agregado la segunda vez que fueron evaluades. Este aumento no se debió solamente a un pequeño incremento en la precisión media, sino que estuvo dado principalmente por menores tiempos de respuesta en la instancia de reevaluación, que a su vez derivaron en una mayor proporción de ejercicios finalizados. Sin embargo, esta mejoría no fue significativamente mayor para les tutores que para les controles, y por lo tanto no presenta evidencia a favor del efecto de enseñar como facilitador del aprendizaje.

Es posible que el efecto de enseñar sobre el propio aprendizaje no fuese inmediato, sino a largo plazo (Ellis et al., 1998; Fiorella & Mayer, 2013; Hermida et al., 2021; Roscoe & Chi, 2007; Sabol & Wisher, 2001; Semb et al., 1993). Por lo tanto, evaluamos nuevamente a todos los grupos dos años más tarde. Sin embargo, aunque la mayoría de les niñes reportaron recordar el juego, no hubo diferencias de desempeño entre grupos, y fue indistinguible, a su vez, del de niñes que no habían recibido ningún tipo de instrucción.

Con respecto a la falta de evidencia a favor del efecto de enseñar como facilitador del aprendizaje, se ha sugerido que, a través de la explicación y el cuestionamiento, enseñar a une par ayuda a les tutores a reflexionar metacognitivamente acerca de su propia pericia y comprensión (Roscoe & Chi, 2007). Sin embargo, aunque esta mejor comprensión acerca de su propio conocimiento podría socavar la confianza de les enseñantes, como se discute más adelante, podría no traducirse inmediatamente

en beneficios para el aprendizaje, hasta tanto hayan tenido la oportunidad de actualizar su conocimiento. Por lo tanto, si bien les tutores en este estudio podrían haber identificado falencias en su conocimiento cuando enseñaban, es posible que no hayan tenido la oportunidad de corregirlas antes de la reevaluación. La inclusión de una instancia adicional de revisión entre las instancias de *Enseñanza entre pares* y de *Reevaluación* podría ayudar a contrastar esta hipótesis en estudios futuros.

3.2.3.3 Enseñanza entre pares y metacognición

Para aproximarnos al estudio de la relación entre la enseñanza entre pares y la metacognición, analizamos el impacto de las distintas condiciones experimentales sobre una variable relacionada: la confianza sobre el propio desempeño. Por un lado, observamos que mientras los niveles de confianza del grupo control aumentaron en la reevaluación –acompañando los mayores niveles de desempeño en esta etapa–, éstos no cambiaron significativamente en el grupo tutor –a pesar del mismo aumento en el desempeño. Además, en la etapa de *Seguimiento*, dos años después, si bien tanto tutores como aprendices redujeron sus niveles de confianza respecto de la última vez que fueron evaluades en el estudio principal –en sintonía con el menor desempeño generalizado en esta etapa– el grupo control no mostró reducción significativa en estos niveles.

Estos resultados parecen señalar que la enseñanza entre pares habría reducido la confianza acerca del propio desempeño, tanto de tutores como de aprendices, en comparación con aprender de une adulte seguido de repasar a solas. Descontextualizados, estos resultados parecen desalentadores. Sin embargo, sabemos que les niñes suelen mostrar niveles de confianza excesivos (Finn & Metcalfe, 2014). En este contexto, entonces, nuestros resultados podrían indicar que la enseñanza entre pares podría haber modulado las habilidades metacognitivas de les participantes, atenuando así sus niveles de confianza.

Sin embargo, para poder hablar de metacognición deberíamos relacionar confianza con desempeño. Pero la capacidad de nuestros datos es limitada, ya que el estudio no fue diseñado con este fin. De hecho, la instancia de la que obtenemos valores de confianza fue introducida simplemente para detener el avance irreflexivo de les participantes a través de los ejercicios, como se indicó en la sección “3.2.1.3 Interfaz

de usuarie”. Por lo tanto, estudios futuros deberían llevarse a cabo para confirmar la interpretación sugerida (ver sección “5.4.3 Metacognición” del Capítulo 5).

3.2.3.4 Diferencias entre géneros

Otro aspecto que analizamos en este estudio fue el de posibles diferencias entre géneros. Muchos estudios han descripto diferencias en desempeño e interés en dominios técnicos entre géneros (Edwards et al., 1997; Hill et al., 2010; Nourbakhsh et al., 2004). Sin embargo, la mayoría de estas diferencias –sino todas– serían en realidad consecuencia de convenciones y estereotipos sociales (Beilock et al., 2010; Dar-Nimrod & Heine, 2006). Por lo tanto, en este estudio era posible que les niñes participantes estuvieran libres de estos sesgos, debido a un adoctrinamiento cultural posiblemente limitado, producto de su corta edad. No obstante, las niñas participantes mostraron menor puntaje agregado que los niños.

Sullivan y Bers (2013) encontraron resultados similares en un programa de robótica con niñes de nivel inicial (jardín de infantes). Llamativamente, un análisis más detallado de nuestros resultados muestra que su puntaje agregado fue menor no porque fueran menos precisas, sino porque se tomaban más tiempo y por consiguiente finalizaban menos ejercicios. Este hallazgo es compatible con resultados previos que muestran que niñas eran menos propensas a tomar riesgos para alcanzar un objetivo cuando programaban con LOGO que sus compañeros varones, mostrando un comportamiento más cuidadoso o precavido (Yelland, 1993). Esto podría estar relacionado con diferencias de autorregulación entre géneros (McClelland et al., 2015). Estas diferencias, a su vez, podrían ser en parte un efecto de los estereotipos de género (Halpern, 2012; Rippon, 2019), que de acuerdo con algunos estudios podrían manifestarse tan temprano como a partir de los seis años de edad (Bian et al., 2017).

3.3 Conclusiones

En el estudio descripto en este capítulo ampliamos el trabajo desarrollado en el capítulo anterior hacia otro dominio del conocimiento: el de la programación de computadoras. A diferencia de aquél, en éste empleamos un paradigma experimental de enseñanza entre pares típico, con roles de tutore y aprendiz bien definidos, y una tarea lo suficientemente compleja para evitar los efectos techo de

desempeño discutidos en el capítulo anterior. De esta manera, observamos que les niñes pudieron, en efecto, aprender a programar de sus pares, con desempeño en principio indistinguible del de les que aprenden de une adulte; según métodos de evaluación que combinan precisión y tiempo de respuesta, como los que suelen usarse en pruebas escolares. Sin embargo, el análisis detallado reveló que detrás de esta aparente igualdad se esconden diferencias en el equilibrio entre velocidad y precisión, con les niñes que aprendieron de une par adoptando un equilibrio menos conservador, con tiempos de respuesta más cortos, a expensas de una menor precisión. En este capítulo, además, avanzamos sobre la relación entre la enseñanza entre pares y otras habilidades cognitivas, como introdujimos en el Capítulo 1. En particular, acerca del efecto de la enseñanza entre pares sobre la metacognición, evaluada indirectamente aquí a través de la confianza sobre el propio desempeño.

Los alcances de este estudio pueden considerarse en tres ámbitos: el científico, el pedagógico y el social. En el ámbito científico, cabe destacar que sería el primer estudio experimental en abordar la enseñanza entre pares de programación en niñes. Además, ofrece una formalización de la plataforma de programación empleada (es decir, su definición en términos de las ciencias de la computación), sentando bases para su reutilización en otros estudios. En el ámbito pedagógico, en un contexto de creciente interés por incorporar la enseñanza de programación en las escuelas, el estudio provee resultados empíricos a favor del uso de la enseñanza entre pares como medio complementario para este fin, ampliando la evidencia de cómo esta práctica puede ayudar a promover el aprendizaje en dominios complejos, y sugiriendo que es posible aprovechar simultáneamente los beneficios intrínsecos de ambos mundos: el de aprender a programar, y el de la enseñanza entre pares. Además, abre el debate en torno al equilibrio entre precisión y exploración deseado para este dominio, y cómo conseguirlo combinando distintas herramientas pedagógicas. Finalmente, en el ámbito social, en un mundo en el que las computadoras y sus redes se encuentran en las bases de nuestra sociedad global, el estudio presenta a la enseñanza entre pares como factor de relevancia para ampliar la distribución del conocimiento acerca de cómo operan estos fundamentos y de cómo moldearlos, esencial para los valores de democracia y pluralidad, como una

forma de liberar a las generaciones futuras de los suelos infértiles de la dependencia tecnológica, y de guiarlas en cambio hacia la soberanía tecnológica.

Más allá de las mejoras respecto del estudio presentado en el Capítulo 2, y de sus fortalezas, el estudio descripto en este capítulo tiene también sus limitaciones. Por un lado, si bien nuevamente no encontramos que enseñar a une par tuviera consecuencias negativas, tampoco encontramos los beneficios esperados en el desempeño, identificando como causa posible a la falta de una instancia de repaso adicional, previa a la etapa de reevaluación, que permita capitalizar los efectos sobre la confianza acerca del propio conocimiento. Por otra parte, la naturaleza del paradigma experimental, con dos instancias de evaluación para controles y tutores, y sólo una para aprendices, limitó la interpretación de algunos resultados. Asimismo, si bien en este estudio avanzamos sobre aspectos de la tesis omitidos en el Capítulo 2, como ser la relación con otras habilidades cognitivas como la metacognición, su abordaje fue aproximado, y quedan cuestiones sin considerar.

Por lo tanto, en el capítulo siguiente presentaremos dos estudios en los que usamos un paradigma experimental que minimiza la complejidad de distribución y análisis, permitiendo así enfocarnos en la resolución de estas limitaciones, incluyendo la ampliación de la base de participantes y su rango etario, la inclusión de instancias de revisión opcionales durante las evaluaciones, la reevaluación de todos los grupos, y la evaluación directa de la metacognición, como moduladora de la enseñanza entre pares.

Capítulo 4: Enseñanza entre pares de reglas jerárquicas

En los capítulos precedentes se abordó la enseñanza entre pares como estrategia pedagógica para la promoción del aprendizaje en dominios donde éste presenta dificultades: el de los cambios conceptuales y el de la programación de computadoras. En esos capítulos, además, se estudió la relación de la enseñanza entre pares con aspectos demográficos, como el género, y con otras habilidades cognitivas, como la metacognición. Sin embargo, la complejidad en la administración y análisis de las evaluaciones, tanto en el Capítulo 2 como –aunque en menor medida– en el Capítulo 3, limitó el número de participantes y las variables consideradas.

En este capítulo utilizamos el juego de cartas de percepción visual *Set* como tarea, cuya evaluación es relativamente sencilla y automatizable, permitiendo así la incorporación de más participantes, más instancias de evaluación por participante, y la evaluación de otras habilidades cognitivas. Esta simplicidad en el análisis, no obstante, no se corresponde con la complejidad en la explicación de la tarea, haciéndola adecuada para un paradigma de enseñanza entre pares.

Además, en este capítulo se incluyen instancias de revisión luego de la enseñanza entre pares, que podrían ser necesarias –como se comentó en el Capítulo 3– para hacer efectivos los potenciales beneficios de enseñar sobre el propio desempeño. Asimismo, se profundiza el estudio de la relación entre metacognición y enseñanza entre pares, iniciado en el capítulo anterior.

Presentaremos dos estudios: el primero, en el que identificamos factores que modulan el desempeño en la tarea; y el segundo, en el que aplicamos el paradigma para el estudio de la enseñanza entre pares, usando los resultados del primero para ajustar el protocolo utilizado.

4.1 Introducción

En los Capítulos 2 y 3 abordamos la cuestión de la enseñanza entre pares en el contexto de dominios del conocimiento que suelen presentar dificultades para su

aprendizaje, postulando a este tipo de enseñanza como una estrategia pedagógica complementaria para facilitarlo. Sin embargo, la complejidad de administración y análisis de las evaluaciones de desempeño correspondientes limitaron el número de participantes y de variables que fue posible considerar.

Puntualmente, en el estudio principal del Capítulo 2, la evaluación era una entrevista *une-a-une* guionada. Por tal motivo, la administración era laboriosa, pues le entrevistadore debía seguir el guión estrictamente, y les participantes debían ser entrevistades une a la vez. Por otra parte, el análisis de las entrevistas era complejo también, requiriendo el trabajo de une técnique a cargo de la codificación de las respuestas, según un extenso manual de codificación.

En cuanto al estudio del Capítulo 3, la evaluación consistía en una serie de ejercicios de programación de computadora. En este caso, si bien la administración prescindía en mayor parte de la asistencia de le investigadore, no fue posible evaluar varies participantes a la vez tampoco, ya que el *software* y la configuración necesarios eran demasiado complejos para ser instalados en los dispositivos que la institución podía proveer. Por otra parte, en cuanto al análisis de las evaluaciones, si bien finalmente dimos con un método automatizado para procesar las respuestas de les participantes, evaluar el desempeño a partir de éstas no es algo trivial. En nuestro caso, decidimos considerar los tiempos de respuesta y las distancias de edición (entre las respuestas de les participantes y las posibles respuestas correctas), pero bien podría haberse empleado una estrategia diferente.

Por lo tanto, para continuar con nuestros estudios, resultaba necesario emplear una tarea a la vez sencilla de administrar y evaluar, que no obstante fuera lo suficientemente compleja para representar un desafío de enseñanza y aprendizaje. Decidimos emplear una adaptación del juego de cartas de percepción visual *Set*, desarrollado en 1974 por la genetista de poblaciones Marsha Falco, basándonos en nuestra experiencia previa con la tarea. En *Set*, cada carta contiene entre una y tres figuras, y es caracterizada por cuatro atributos: forma, color, textura y número de dichas figuras (Figura 4.1). A su vez, cada atributo puede tomar uno de tres valores posibles: diamante, óvalo o garabato (forma); rojo, azul o verde (color); lleno, rayado o vacío (textura); y uno, dos o tres (número).

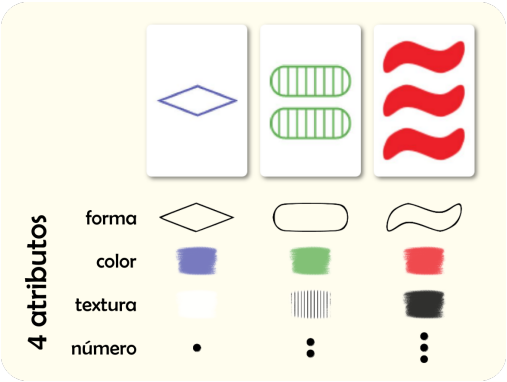


Figura 4.1: *Atributos y valores posibles*

Cada carta es caracterizada por cuatro atributos: forma, color, textura y número de las figuras que contiene. A su vez, cada uno de estos atributos puede tomar uno de tres valores posibles.

El objetivo del juego es identificar cuáles tres cartas forman un Set. Los Sets se describen en dos niveles de abstracción, por lo que les participantes deben aplicar un sistema de reglas jerárquicas para identificarlos (Eckstein et al., 2019). Por un lado, la descripción de bajo nivel comprende patrones dentro de cada atributo, considerado independientemente de los demás. Concretamente, en un Set cada atributo (forma, color, textura y relleno) debe adoptar uno de dos patrones válidos posibles: patrón *igual*, en que todas las cartas son iguales, o patrón *diferente*, en que todas son distintas, siempre para el atributo en consideración. Nos referiremos a esta regla de bajo nivel como “regla Set”. En cambio, diremos que un atributo adopta un patrón inválido *roto* si dos cartas son iguales y una distinta para un atributo dado. Por otra parte, la descripción de alto nivel comprende las posibles combinaciones de patrones entre los distintos atributos. Puntualmente, en un Set los atributos deben adoptar alguna combinación de patrones válidos *igual* y *diferente*. Así, los Sets se subclasifican según el número de atributos que adoptan el patrón *diferente* (ver sección “4.2.1.2 Clasificación de ejercicios”). En cambio, si alguno de los atributos presenta un patrón *roto*, entonces las tres cartas no forman un Set (NoSet). Ver ejemplos en Figura 4.2.

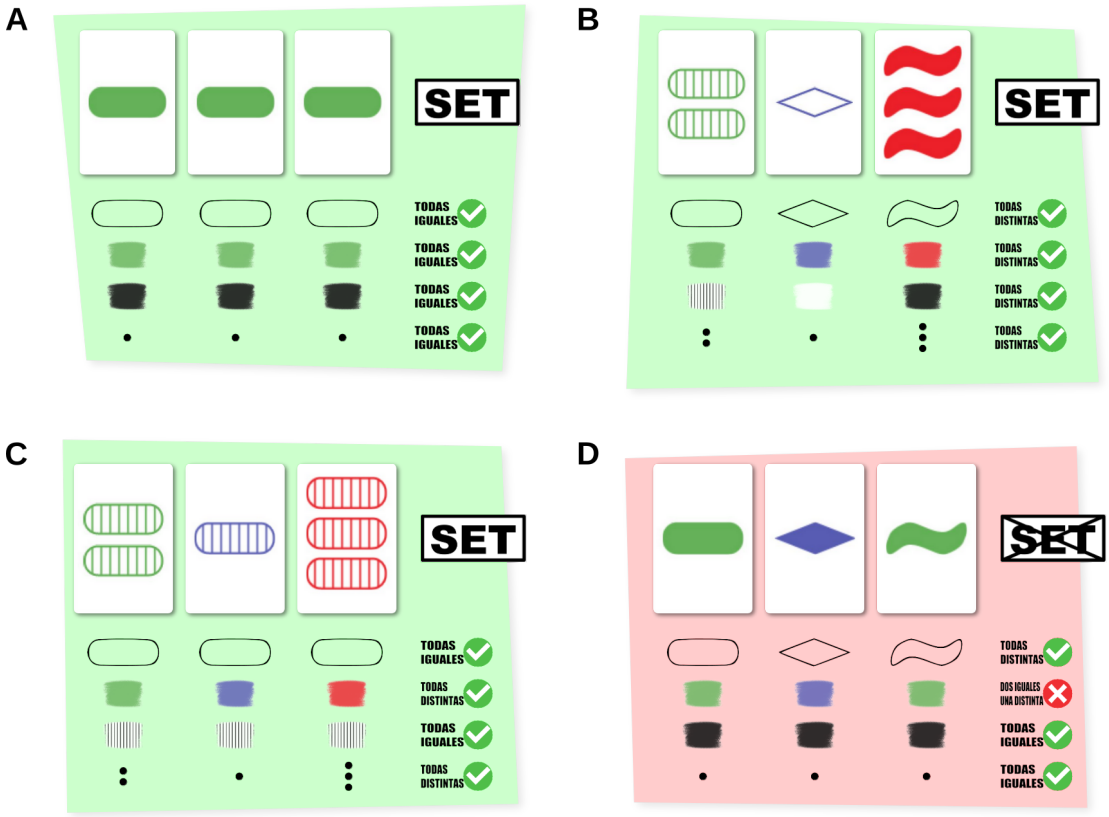


Figura 4.2: *Ejemplos de Set y NoSet*

(A) **Ejemplo de Set 0-diferente.** La regla Set se cumple para el atributo forma pues las cartas son *todas iguales* (óvalo) para este atributo (patrón *igual*). Lo mismo ocurre para los demás atributos. Como la regla Set se cumple para todos los atributos, la regla de alto nivel queda satisfecha, y las tres cartas forman un Set. Cabe aclarar que esta subclase de Set 0-diferente es hipotética, ya que no es posible tener dos o más cartas idénticas; se incluye aquí con fines explicativos.

(B) **Ejemplo de Set 4-diferente.** La regla Set se cumple para el atributo color, pues las cartas son *todas distintas* (verde, azul y rojo) para este atributo (patrón *diferente*). Lo mismo ocurre para los demás atributos. Como la regla Set se cumple para todos los atributos, las tres cartas forman un Set.

(C) **Ejemplo de Set 2-diferente.** La regla Set se cumple para el atributo textura, pues las cartas son *todas iguales* (rayado) para este atributo (patrón *igual*). Lo mismo ocurre para el atributo forma. Por otra parte, la regla Set también se cumple para el atributo número, pues las cartas son *todas distintas* (uno, dos y tres) para este atributo (patrón *diferente*). Lo mismo ocurre para el atributo color. Como la regla Set se cumple para todos los atributos, las tres cartas forman un Set.

(D) **Ejemplo de NoSet 1-roto.** La regla Set no se cumple para el atributo color, pues las cartas no son ni *todas iguales*, ni *todas distintas* para este atributo (patrón *roto*); son *dos iguales* (verde) y *una distinta* (azul). Como la regla Set no se cumple para al menos un atributo, la regla de alto nivel no se satisface, y las tres cartas no forman un Set (NoSet).

En la versión original del juego, se disponen varias cartas sobre una mesa, y les jugadores deben identificar Sets entre ellas. En nuestras adaptaciones, en cambio, se muestran tres cartas por vez, y les participantes deben decidir si forman un Set o no. De este modo, la tarea reúne las características que estábamos buscando. Por un lado, su administración es simple, automatizable y adaptable a plataformas estándar (como veremos más adelante), permitiendo administrarse simultáneamente a varies participantes a la vez, con un mínimo de intervención por parte de le investigadore. Por otra parte, la evaluación del aprendizaje en términos del desempeño es sencilla y concreta, reduciéndose a considerar la proporción de respuestas correctas sin ambigüedad, por la naturaleza binaria de las mismas (Set o NoSet), y los tiempos de respuesta.

Estas condiciones permiten ampliar (i) el número de participantes, extendiendo el rango etario considerado; (ii) el número de evaluaciones por participante, permitiendo comparar la enseñanza entre pares y la enseñanza tradicional no sólo en la instancia de evaluación, sino en la de reevaluación también; y (iii) el número de variables evaluadas, posibilitando la profundización del estudio de la relación entre metacognición y enseñanza entre pares, iniciada en el Capítulo 3. Estas ventajas se dan sin detrimento de la idoneidad de este paradigma para estudiar la enseñanza entre pares, ya que la estructura jerárquica de sus reglas, en apariencia simples, implica un ejercicio no trivial de abstracción y ejemplificación para lograr transmitirlas satisfactoriamente.

Si bien la estructura de *Set* ha motivado varios estudios en el campo de las matemáticas (Davis & Maclagan, 2003), sus características cognitivas y perceptuales no han sido estudiadas en la misma profundidad (Eckstein et al., 2019; Jacob & Hochstein, 2008). Por tanto, en primer lugar, en el *Estudio inicial* a continuación identificamos factores perceptuales y cognitivos que modulan el desempeño en la tarea.

Luego, en el *Estudio principal*, empleamos los factores identificados en el *Estudio inicial* para ajustar el paradigma al estudio de la enseñanza entre pares, ampliando así el reconocido potencial de este paradigma como herramienta para estudiar preguntas en el campo de la cognición (Yuviler-Gavish et al., 2020).

4.2 Estudio inicial

En este trabajo examinamos factores perceptuales y cognitivos que modulan el desempeño en el juego *Set*,[28] con el objetivo de caracterizar este paradigma para su uso en el *Estudio principal* sobre enseñanza entre pares. En particular, comparamos (i) Sets vs NoSets, (ii) Sets en función del número de atributos con patrón *diferente*, (iii) NoSets en función del número de atributos con patrón *roto*, y (iv) la saliencia perceptual de los distintos atributos.

4.2.1 Métodos

4.2.1.1 Participantes y procedimiento

30 estudiantes de un curso universitario de neurociencia cognitiva en la ciudad de Buenos Aires participaron del estudio. Les participantes dieron consentimiento vía correo electrónico para la publicación de los resultados obtenidos.

Les participantes tomaban una sesión frente a una computadora de escritorio. La sesión comprendía de dos partes, una de *práctica* y otra de *evaluación*. La parte de práctica incluía un bloque de diez ejercicios, y la de evaluación 3 bloques de 40 ejercicios cada uno (ver Tabla 4.1). Les participantes podían decidir libremente cuánto tiempo tomarse entre bloque y bloque. Antes de comenzar, recibían una explicación autoguiada en pantalla con las reglas del juego y dos ejercicios de ejemplo: un Set 2-diferente, y un NoSet.

En cada ejercicio se mostraban tres cartas y les participantes debían responder con el teclado si éstas formaban un Set o no. El tiempo máximo por ejercicio era de 10 segundos. Ejercicios sin responder fueron tomados como incorrectos para el

28 Este trabajo fue realizado con la colaboración de Nicolás Roulet y Anahí Rosso, como parte del curso de posgrado Seminario de Neurociencia Computacional. Los resultados fueron presentados en la XVII Reunión Nacional y VI Encuentro Internacional de la Asociación Argentina de Ciencias del Comportamiento, en Posadas, Argentina. https://doi.org/10.6084/m9.figshare.13103534.v1

análisis. En caso de responder incorrectamente, se mostraba una pantalla roja brevemente antes de continuar con el ejercicio siguiente.

Sólo los ejercicios de los bloques de evaluación fueron considerados para calcular el desempeño (precisión y tiempos de respuesta). Les participantes con precisión menor al 80% (n=4) fueron excluidos de la muestra.

4.2.1.2 Clasificación de ejercicios

Los ejercicios se clasifican según tres niveles de clasificación (ver Tabla 4.1):

Clase. En primer lugar, los ejercicios se clasifican en Set y NoSet, según los atributos de las cartas adopten alguna combinación de patrones válidos *igual* y *diferente*, como se indicó en la Introducción.

Subclase. A su vez, los Sets se subclasifican según la cantidad de atributos con patrón *diferente* (es decir, para los cuales las cartas son distintas entre sí); de 1-diferente a 4-diferente. No existen ejercicios de la subclase Set 0-diferente pues implicaría que las tres cartas fuesen idénticas. Por su parte, los NoSets se subclasifican según la cantidad de atributos con patrón *roto* (es decir, para los cuales dos cartas son iguales y una distinta); de 1-roto a 4-roto. No existen ejercicios de la subclase NoSet 0-roto, pues serían un Set.

Tipo. Finalmente, los Sets se tipifican en función de la combinación específica de atributos con patrón *igual*/*diferente*. Así, hay 4 tipos de Set 1-diferente, 6 tipos de Set 2-diferentes, 4 tipos de Set 3-diferentes, y 1 tipo de Set 4-diferente. Lo mismo ocurre para los NoSets, según la distribución específica de atributos con patrón *roto*. En particular, en las subclases Set 1- y 3-diferente y NoSet 1- y 3- roto, los tipos se identifican según el atributo extraño, es decir aquél que sigue un patrón distinto de los otros tres. Es decir, la subclase Set 1-diferente se identifica según el atributo con patrón *diferente*, la subclase Set 3-diferente se identifica según el atributo con patrón *igual*, la subclase NoSet 1-roto se identifica según el atributo con patrón *roto*, y la subclase NoSet 3-roto según el atributo con patrón válido (no *roto*). Ver Tabla 4.1.

Clase	Subclase	Tipo				Atributo	Estudio incial		Estudio principal		Recuento (*)		
		F	C	T	N	extraño	Práctica	Eval.	Práctica	Eval.			
Set	0-diferente	=	=	=	=	-	0	0	1	0	81	81	81**
	1-diferente	≠	=	=	=	F	2†	6	0	1	27	108	1080
		=	≠	=	=	C		6		1	27		
		=	=	≠	=	T		6		1	27		
		=	=	=	≠	N		6		1	27		
	2-diferente	≠	≠	=	=	-	0	2	0	0	54	324	
		≠	=	≠	=	-		2			54		
		≠	=	=	≠	-		2			54		
		=	≠	=	≠	-		2			54		
		=	=	≠	≠	-		2			54		
		=	≠	≠	=	-		2			54		
	3-diferente	=	≠	≠	≠	F	2†	6	0	1	108	432	
		≠	=	≠	≠	C		6		1	108		
		≠	≠	=	≠	T		6		1	108		
		≠	≠	≠	=	N		6		1	108		
	4-diferente	≠	≠	≠	≠	-	1	6	0	4	216	216	
NoSet	1-roto	✗	✓	✓	✓	F			0	1	2106	8424	84240
		✓	✗	✓	✓	C			1	1	2106		
		✓	✓	✗	✓	T			0	1	2106		
		✓	✓	✓	✗	N			0	1	2106		
	2-roto	✗	✗	✓	✓	-	5‡	54‡	0	0	4212	25272	
		✗	✓	✗	✓	-					4212		
		✗	✓	✓	✗	-					4212		
		✓	✗	✓	✗	-					4212		
		✓	✓	✗	✗	-					4212		
		✓	✗	✗	✓	-					4212		
	3-roto	✓	✗	✗	✗	F			0	1	8424	33696	
		✗	✓	✗	✗	C				1	8424		
		✗	✗	✓	✗	T				1	8424		
		✗	✗	✗	✓	N				1	8424		
	4-roto	✗	✗	✗	✗	-			0	0	16848	16848	
						Totales	**10**	**120**	**2**	**20**			**85320**

Tabla 4.1: *Clasificación de ejercicios*

† El atributo con patrón *diferente* o *igual* en las subclases Set 1-diferente o 3-diferente, respectivamente, fue seleccionado al azar, pero distinto cada vez (tal que todos los atributos fuesen el atributo extraño en algún ejercicio Set 1-diferente o 3-diferente en la parte de práctica)

‡ La subclase y tipo específico fueron seleccionados al azar, según la proporción relativa indicada en las columnas “Recuento”

* Recuento de los grupos de tres cartas posibles (sin importar el orden de colocación) para cada clase, subclase y tipo.

** La subclase Set 0-diferente es hipotética y no se incluye en el recuento total.

Atributos: F, forma; C, color; T, textura; N, número.

=, igual; ≠, diferente; ✓, válido; ✗, roto.

En la parte de práctica se incluyeron 2 ejercicios Set 1-diferente, 2 Set 3-diferente, 1 Set 4-diferente, y 5 NoSets. En la parte de evaluación se incluyeron 6 ejercicios por tipo de las subclases Set 1-diferente, 3-diferente y 4-diferente, 2 ejercicios por tipo de la subclase Set 2-diferente, y 54 NoSets (ver tabla 4.1). El orden de presentación fue aleatorio.

Existe más de un conjunto de tres cartas para cada tipo específico de ejercicio. Por ejemplo, para un ejercicio de la clase Set, subclase 1-diferente, tipo forma-diferente, existen 27 grupos de tres cartas, sin importar el orden de colocación (ver Tabla 4.1, Recuento). En todos los casos, el grupo específico que se mostró fue seleccionado al azar.

Cabe destacar que la clasificación de NoSets en subtipos y en función del atributo extraño no se consideró hasta luego de recolectados los datos. Por lo tanto, si bien el número de Sets según subclase y tipo fue controlado, este no fue el caso para los NoSets.

4.2.1.3 Análisis

Se verificó el cumplimiento de todos los supuestos de las pruebas estadísticas paramétricas utilizadas; se usaron pruebas no paramétricas donde estos supuestos no se cumplían.

4.2.2 Resultados

4.2.2.1 Precisión

En primer lugar se comparó la proporción de respuestas correctas (precisión) entre ejercicios Set y NoSet. No es posible realizar una prueba *t* de muestras apareadas porque la distribución de las diferencias (precisión en ejercicios Set – precisión en ejercicios NoSet) no es normal, según un test Shapiro-Wilk (W = 0,87, p = 3,05E-03). Por otra parte, tampoco se puede realizar un test no paramétrico de Wilcoxon porque estas diferencias tampoco están simétricamente distribuidas alrededor de la mediana. Por lo tanto, se usó una prueba no paramétrica del signo, según la cual los ejercicios Set fueron respondidos correctamente en mayor proporción que los ejercicios NoSet, M = 8,0, p = 1,54E-03 (Figura 4.3). Esto sugiere que los ejercicios Set podrían ser más fáciles de identificar que los ejercicios NoSet.

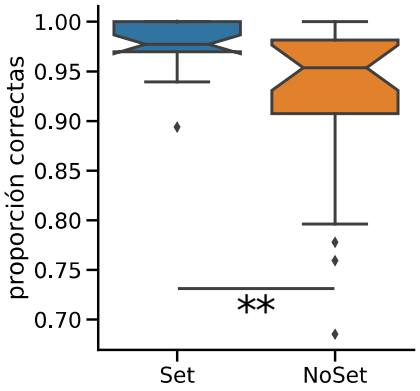


Figura 4.3: *Precisión por clase*

La proporción de respuestas correctas o precisión fue significativamente mayor para ejercicios de la clase Set que para ejercicios de la clase NoSet. **: p < 0,01

A continuación analizamos los tiempos de respuestas, una variable más adecuada para análisis más detallados, dada la naturaleza continua de los valores que puede adoptar.

4.2.2.2 Tiempo de repuesta

A continuación, se compararon los ejercicios en función del tiempo de respuesta. Se contemplaron únicamente los ejercicios respondidos correctamente (Figura 4.4A). Como es habitual en estos casos, se aplicó una transformación logarítmica para obtener una distribución normal de la variable considerada.

En primer lugar, los tiempos de respuesta resultaron significativamente menores en los ejercicios NoSet, IC 95% (seg) [0,44, 0,79], que en los ejercicios Set, IC 95% (seg) [1,13, 1,81], como resulta de una prueba *t* para muestras apareadas, $t = 3{,}26$, $p = 3{,}21E\text{-}03$ (Figura 4.4B).

A continuación, se puso el foco dentro de cada una de las clases (Set y NoSet), y en función del número de atributos con patrón *diferente* o *roto*, respectivamente. En cuanto a los Sets, nuestra predicción era que la “diferencia” sería más difícil de reconocer que la igualdad, y que por lo tanto los Sets con mayor número de atributos con patrón *diferente* serían más difíciles de identificar correctamente. Por otra parte, también esperábamos que los NoSets con mayor número de atributos con patrón *roto* serían identificados más rápidamente, visto que con sólo identificar un atributo “roto” es suficiente para afirmar que el ejercicio es un NoSet.

En ambos casos se realizó una regresión con modelo lineal mixto, con participante (n=26) como variable aleatoria. En efecto, en ejercicios Set correctos (n=1674, 64,4 por participante), el logaritmo del tiempo de respuesta crece con el número de atributos “diferente” a una tasa (7,98E-02) significativamente distinta de cero, z = 86,6, p = 4,66E-77. Es decir, el tiempo de respuesta aumenta 20% por cada atributo “diferente” adicional (Figura 4.4C). Por otra parte, en ejercicios NoSet correctos (n=1293, 49,7 por participante), el logaritmo del tiempo de respuesta decrece con el número de atributos “roto” a una tasa (–2,70E-02) significativamente distinta de cero, z = –4,74, p = 2,10E-06. Es decir, el tiempo de respuesta disminuye 6% por cada atributo “roto” adicional (Figura 4.4D).

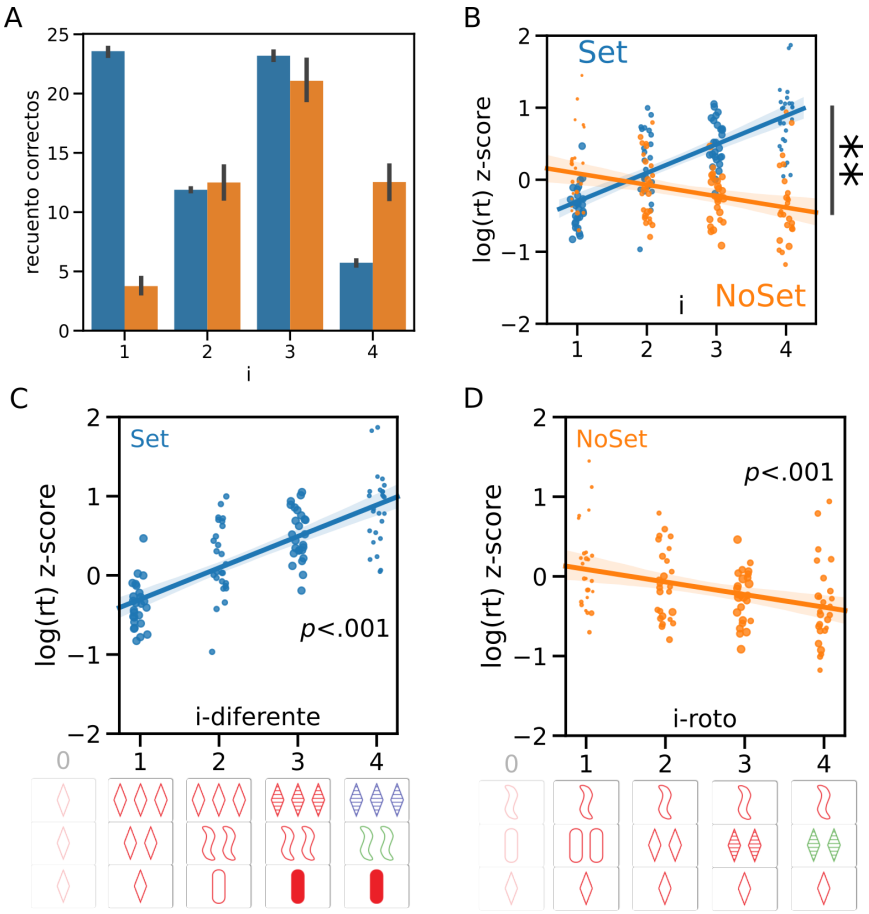


Figura 4.4: *Tiempos de respuesta para ejercicios correctos*

i: número de atributos con patrón *diferente* para ejercicios Set (azul), o *roto* para ejercicios NoSet (naranja). Para facilitar la visualización, ya que hay algunes sujetos que demoran sistemáticamente más, y otros menos, se usa el z-score con respecto al tiempo de respuesta medio por participante.

(A) Número promedio de ejercicios respondidos correctamente por subclase y

participante. La mayor dispersión para ejercicios NoSet posiblemente se deba a que, a diferencia de los ejercicios Set, la proporción entre las distintas subclases no fue controlada (ver sección “4.2.1.2 Clasificación de ejercicios”).

(B) En promedio, la identificación correcta de NoSets es más rápida que la de Sets.

(C) La identificación correcta de Sets es más rápida cuanto menor es el número de atributos con patrón *diferente*. Se incluyen ejemplos ilustrativos de Sets 1- a 4-diferente.

(D) La identificación correcta de NoSets es más rápida cuanto mayor es el número de atributos con patrón *roto*. Se incluyen ejemplos ilustrativos de NoSets 1- a 4-roto.

Las barras de error en la figura A y las bandas translúcidas en las figuras B, C y D representan intervalos de confianza del 95% obtenidos con el método *bootstrap* de remuestreo. En estas últimas figuras, cada punto es un sujeto y el tamaño de los puntos representa el número de muestras (es decir, el número de ejercicios correctos para la subclase respectiva para cada sujeto). **: $p < 0,01$.

4.2.2.3 Saliencia de atributos

Finalmente, a fin de caracterizar la saliencia intrínseca de los distintos atributos, se compararon los tiempos de respuesta para ejercicios Set 1- y 3-diferente y para ejercicios NoSet 3-roto en función del *tipo* específico o atributo extraño (ver sección “4.2.1.2 Clasificación de ejercicios”). Como en la sección precedente, consideramos ejercicios correctos solamente. Los ejercicios NoSet 1-roto no fueron considerados porque el número de ejercicios NoSet 1-roto era muy pequeño (Figura 4.4A), ya que a diferencia de los ejercicios Set la proporción de las distintas subclases de ejercicios NoSet no fue controlada.

Se hicieron pruebas ANOVA de medidas repetidas, una por cada subclase considerada, con atributo extraño como factor intra-sujeto. Se encontraron efectos significativos del atributo extraño sobre los tiempos de respuesta, y se hicieron análisis post-hoc con prueba HSD de Tukey (ver Figura 4.5 y Tabla 4.2).

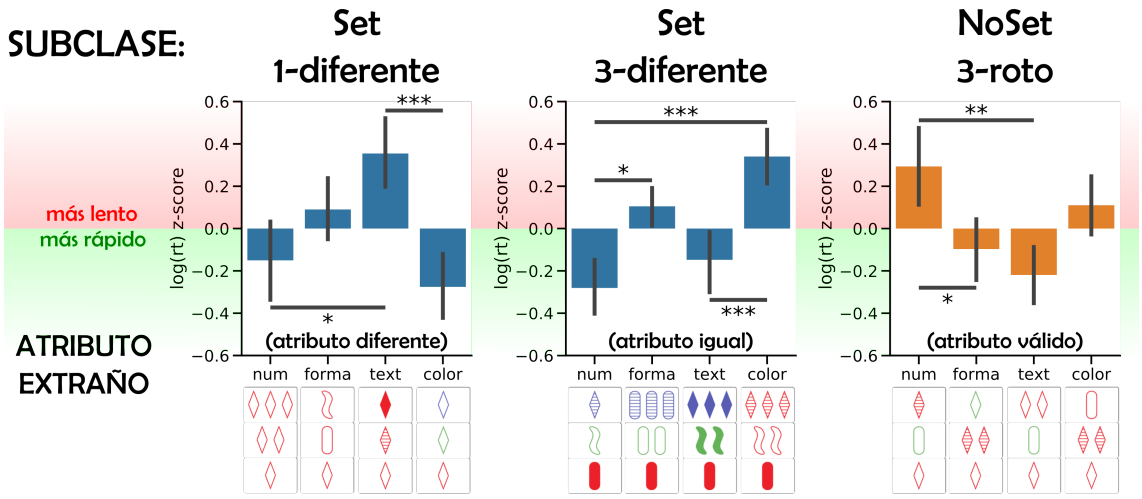


Figura 4.5: *Saliencia de atributos*

Para caracterizar la saliencia de los distintos atributos, se compararon los tiempos de respuesta en función del *tipo* específico de ejercicio o atributo extraño para las respuestas correctas de los ejercicios Set 1-diferente, Set 3-diferente y NoSet 3-roto (se incluyen ejemplos ilustrativos de cada *tipo* debajo de las figuras). A los tiempos de respuesta se les aplica una transformación logarítmica y se expresan como z-score o número de desvíos respecto del tiempo de respuesta medio para cada subclase. Ver Tabla 4.2 para una descripción estadística detallada. Las barras de error representan intervalos de confianza del 95% obtenidos con el método *bootstrap* de remuestreo. *: p < 0,05, **: p < 0,01, ***: p < 0,001.

Clase	Subclase	ANOVA		Tukey (p)		
		F	p	C	N	F
Set	1-diferente	5,92	0,002	**N** 0,673 (ns)		
				F 0,112 (ns)	0,660 (ns)	
				T 0,001 (**)	0,025 (*)	0,305 (ns)
Set	3-diferente	10,25	<0,001	**N** <0,001 (***)		
				F 0,133 (ns	0,04 (*)	
				T <0,001 (***)	0,941 (ns)	0,135 (ns)
NoSet	3-roto	5,52	0,003	**N** 0,680 (ns)		
				F 0,352 (ns)	0,034 (*)	
				T 0,054 (ns)	0,002 (**)	0,785 (ns)

Tabla 4.2: *Pruebas estadísticas para la saliencia de atributos*

Resultados de las pruebas ANOVA para cada subclase considerada, con atributo extraño como factor. Se muestran también los resultados de las comparaciones post-hoc entre los distintos atributos extraños (N: número, F: forma, T: textura, C:

color) con prueba HSD de Tukey, con las correcciones correspondientes por comparaciones múltiples.

4.2.3 Discusión

En este *Estudio inicial* examinamos factores perceptuales y cognitivos que modulan el desempeño en el juego *Set*, con el fin de caracterizar este paradigma para su uso en el *Estudio principal* sobre enseñanza entre pares a continuación.

En primer lugar, se observó que la proporción de respuestas correctas fue mayor para ejercicios Set que para ejercicios NoSet. Este resultado sugiere que los ejercicios Set podrían ser más fáciles de identificar. Sin embargo, no se puede descartar que responder “Set” sea un comportamiento por defecto o aprendido (en especial teniendo en cuenta que hay un 22% más de ejercicios Set que NoSet en la parte de evaluación). Esta posible respuesta predeterminada resultaría en una respuesta correcta en un ejercicio Set, y en una incorrecta en uno NoSet, explicando la asimetría de precisión entre ambas clases de ejercicio.

Por otra parte, los tiempos de respuesta resultaron más cortos para la identificación correcta de NoSets que de Sets. Esto puede ser porque mientras en los ejercicios Set se debe verificar que cada uno de los atributos adopte uno de los dos patrones válidos posibles (*igual* o *diferente*), en los ejercicios NoSet es suficiente con encontrar un atributo que adopte el patrón *roto*. En línea con esta idea, los ejercicios NoSet fueron correctamente identificados más rápidamente a mayor número de atributos “roto”. De hecho, los ejercicios NoSet 1-roto (los más lentos de identificar) parecen demorar el mismo tiempo que los ejercicios Set más fáciles (1-diferente). Asimismo, el tiempo necesario para identificar Sets correctamente también varía en función del número de atributos con patrón *diferente*, siendo los Set 4-diferente los que requieren más tiempo. Esto sugiere que la igualdad sería más evidente que la diferencia.

Estos hallazgos, acerca de las diferencias entre Sets y NoSets, y de la preferencia por la igualdad por sobre la diferencia, son consistentes con estudios previos. Con respecto a las diferencias entre Sets y NoSets, si bien el menor tiempo de respuesta para NoSets que para Sets fue descripto en estos estudios previos (Eckstein et al.,

2019), la mayor precisión para Sets que para NoSets no parece haber sido descripta previamente. Con respecto a la preferencia por la igualdad por sobre la diferencia, nuestros resultados son consistentes con estudios previos que describieron menor precisión (Eckstein et al., 2019), y mayor tiempo de respuesta (Eckstein et al., 2019; Jacob & Hochstein, 2008) para mayor número de atributos con patrón *diferente*. Por un lado, la expresión formal de los atributos "diferente" es más larga que la de los atributos "igual"; es decir, presentan mayor *complejidad booleana* (Eckstein et al., 2019), lo que a su vez resultaría en mayor dificultad subjetiva (Feldman, 2000). Además, para cada atributo "diferente" tres valores distintos deben alojarse en la memoria de trabajo de le participante, mientras que para un atributo "igual" sólo debe alojarse el mismo valor repetido tres veces (Eckstein et al., 2019).

Cabe destacar que esta consistencia con estudios previos se encontró a pesar de ciertas diferencias en la metodología empleada. En particular, en nuestro estudio (i) se presentaron grupos de tres cartas aisladas, para decidir si forman un Set o no, lo cual reporta ciertas ventajas por sobre la forma original del juego (Jacob & Hochstein, 2008) en la que se muestra una matriz de cartas en la que hay que encontrar un Set; y (ii) las cartas se presentaron simultáneamente y no en serie (Eckstein et al., 2019).

Con respecto a los ejercicios NoSet, nuestros resultados muestran que a mayor número de atributos con patrón *roto*, mayor el tiempo de respuesta. Esto es consistente con la idea de que les participantes responden NoSet tan pronto encuentran un atributo roto (sin seguir evaluando los demás atributos). Si bien un aspecto obvio, no parece haber sido reportado anteriormente.

Finalmente, para evaluar una posible saliencia intrínseca de los atributos, consideramos tres subclases de ejercicios, alterando el atributo extraño con el objetivo de minimizar los tiempos de respuesta. En los ejercicios Set 1-diferente, sería preferible que el único atributo con patrón *diferente* fuese de saliencia elevada, para compensar la baja saliencia de la diferencia (frente a la igualdad) descripta previamente. Contrariamente, en los ejercicios Set 3-diferente, sería preferible que el único atributo con patrón *igual* fuese el de menor saliencia, para compensar su baja saliencia con la alta saliencia de la igualdad. Finalmente, en los ejercicios NoSet 3-roto, sería preferible que el único atributo válido (no *roto*) fuese de baja saliencia, para aliviar el conflicto que podrían suponer este tipo de atributos en un NoSet. De

acuerdo con los resultados y según las preferencias mencionadas, el atributo color se manifestó sistemáticamente como un atributo de alta saliencia, mientras que el atributo textura se manifestó sistemáticamente como atributo de baja saliencia. Los atributos número y forma variaron en su saliencia en función de la subclase de ejercicio considerada. Si bien un estudio previo indagó en la saliencia de atributos, sólo encontraron preferencias consistentes a nivel individual (Eckstein et al., 2019).

Cabe señalar que si bien en este estudio consideramos las variables *precisión* y *tiempo de respuesta* como medidas separadas del desempeño, un análisis alternativo podría haber considerado la variable *eficiencia inversa*, que divide el tiempo de respuesta por la precisión (Townsend & Ashby, 1983). Sin embargo, su variabilidad suele ser mayor por combinar dos variables sujetas a error de muestreo, no es claro que dividir tiempo de respuesta por precisión refleje adecuadamente el peso relativo de ambas, y en general su consideración conjunta no siempre es beneficiosa (Bruyer & Brysbaert, 2011).

Los resultados de este estudio, acerca de los factores que modulan el desempeño en esta tarea, fueron utilizados para ajustar el protocolo del *Estudio principal* detallado a continuación (específicamente, qué tipos de ejercicios presentar y en qué orden), en el que aplicamos este paradigma para el estudio de la enseñanza entre pares.

4.3 Estudio principal 3

En este estudio profundizamos nuestra investigación sobre enseñanza entre pares, aprovechando las ventajas del paradigma explorado en el *Estudio inicial*. Específicamente, la facilidad de administración y evaluación de este paradigma nos permitió incluir: (1) un mayor número de participantes, ampliando el rango etario considerado; (2) un mayor número de evaluaciones por participante, optimizando las comparaciones entre grupos; y (3) un mayor número de variables, profundizando la exploración de la interacción con otras habilidades cognitivas, específicamente la metacognición.

Brevemente, primero un grupo de niñes fueron enseñades por une adulte. Luego, una parte de este grupo (tutores) tuvo que enseñar a une compañere (aprendiz), mientras que la otra parte (controles) tuvo que repasar a solas. Todes les niñes

fueron evaluades dos veces: después de aprender de une adulte o par (evaluación), y luego de enseñar o repasar (reevaluación). Durante las evaluaciones se mostraban 20 ejercicios con tres cartas cada uno y les niñes debían decidir si formaban un Set o no. Los detalles del paradigma experimental fueron ajustados teniendo en cuenta la caracterización realizada en el Estudio inicial de este capítulo. Además, antes de mostrarles la respuesta correcta, debían decidir si deseaban aceptar o rechazar una explicación del ejercicio, como indicador indirecto de la confianza sobre el propio desempeño.

El objetivo de este estudio es por tanto evaluar cómo la enseñanza entre pares en niñes impacta en el aprendizaje, si depende de habilidades metacognitivas, y cómo estos efectos cambian con la edad. Específicamente, nos enfocamos en: (1) desempeño, comparando (a) aprender de une adulte, vs (b) aprender de une par (tipo de aprendizaje); (2) desempeño, comparando (a) enseñar a une par, vs (b) repasar a solas (factor tutor); (3) efectos de (i) tipo de aprendizaje, y (ii) factor tutor, sobre la confianza acerca del propio conocimiento; y (4) la edad, el género, la confianza y la metacognición como moduladores de (1) y/o (2).

4.3.1 Métodos

4.3.1.1 Resumen

Para este estudio realizamos una nueva adaptación para computadora del juego de cartas *Set*. Para una descripción general del juego, ver la sección “4.1 Introducción” de este capítulo; para una descripción de la adaptación particular realizada para este estudio, ver la sección “4.3.1.5 Plataforma experimental”, a continuación.

En resumen, todes les niñes participaron de cuatro etapas; ver sección “4.3.1.6 Etapas” para una descripción detallada. Primero, en la etapa de *Instrucción* fueron enseñades por une adulte experte (condición *adulto-instructor*) o por une par (condición *par-instructor* = grupo *aprendiz).* Luego, en la etapa de *Evaluación* fueron evaluades por primera vez. A continuación, en la etapa de *Enseñanza/Revisión* algunes debieron enseñar a une par (grupo *tutor*), mientras que otres debieron repasar a solas (grupos *control* y *aprendiz*). Finalmente, en la etapa de *Reevaluación* les participantes fueron evaluades por segunda vez. Ver secciones

"4.3.1.3 División por grupos y parejas" y "4.3.1.6 Etapas" para descripciones detalladas de estos aspectos.

4.3.1.2 Muestra

Participaron 126 niñes de 2do, 4to y 6to grado (3 clases por grado) de una escuela primaria bilingüe de nivel socioeconómico medio-alto y alto, en el Área Metropolitana de Buenos Aires, Argentina. Les padres o tutores de todes les participantes firmaron voluntariamente el consentimiento previamente autorizado por el Comité de Ética de la Dirección de Investigación del Centro de Educación Médica e Investigación Clínica "Norberto Quirno" (CEMIC), Unidad Asociada del CONICET, Protocolo Nº 683.

Los datos de dos participantes fueron descartados pues habían participado previamente de un estudio preliminar; para no excluir a estes participantes de la actividad, fueron asignados al grupo control para minimizar el impacto de su experiencia previa sobre les demás participantes. Ver Tabla 4.3 para la distribución por grado, clase, género (según fue informado por las autoridades escolares) y condición/grupo experimental, y Tabla 4.4 para la distribución por edades.

Grado	Clase	Subcl.	Adulto-instructor				Par-instructor		< *Condición*
			Control		Tutor		Aprendiz		< *Grupo*
			F	M	F	M	F	M	< *Género*
2	2A	-	2	2	2	2	2	2	**12**
	2B	-	3	1	2	2	2	2	**12**
	2C	-	2	-	2	3	2	3	**12**
		Subtotal	**7**	**3**	**6**	**7**	**6**	**7**	**36**
		Total	**10**		**13**		**13**		
4	4A	**1**	4	1	1	1	1	1	**18**
		2	1	2	2	1	2	1	
	4B	-	3	2	3	3	3	3	**17**
	4C	**1**	1	2	2	2	2	2	**22**
		2	2	3	2	1	2	1	
		Subtotal	**11**	**10**	**10**	**8**	**10**	**8**	**57**
		Total	**21**		**18**		**18**		
6	6A	-	3	2	3	1	3	1	**13**
	6B	-	-	1	1	2	1	2	**7**
	6C	-	3	-	3	1	3	1	**11**
		Subtotal	**6**	**3**	**7**	**4**	**7**	**4**	**31**
		Total	**9**		**11**		**11**		

Tabla 4.3: *Distribución por grado, clase, género y condición/grupo*

Para la distribución en grupos, el objetivo fue que el número de participantes por grupo y género (F: femenino, M: masculino) estuviera balanceado en cada grado.

Grado	Grupo	Edad (años, meses)		
		Mín	**Máx**	**Media**
2	**C**	6, 8	7, 4	6, 12
	A	6, 10	7, 8	7, 3
	T	6, 8	7, 7	7, 2
	Gral	6, 8	7, 8	7, 2
4	**C**	9, 4	10, 3	9, 11
	A	8, 8	9, 8	9, 2
	T	8, 8	10, 3	9, 9
	Gral	8, 8	10, 3	9, 8
6	**C**	11, 4	12, 1	11, 8
	A	11, 5	12, 3	11, 11
	T	11, 3	12, 2	11, 9
	Gral	11, 3	12, 3	11, 10

Tabla 4.4: *Edad de les participantes*

Edad mínima, máxima y media por grado y grupo experimental (C: control, A: aprendiz, T: tutor). Gral: edad mínima, máxima y media por grado, sin dividir por grupo experimental. Esta tabla no incluye el dato de dos participantes (une de 2do grado y otre de 4to grado) cuya fecha de nacimiento no pudo ser obtenida.

Además, análisis preliminares revelaron que algunes participantes respondían rápidamente, en menos de 1 seg, aparentemente al azar. Estos tiempos de respuesta se encuentran alejados de los tiempos de respuesta medios, y por lo tanto obstaculizan el uso de pruebas paramétricas. Por consiguiente, optamos por la opción conservadora de excluir a les participantes que tuvieran al menos una sesión (de evaluación o reevaluación; ver sección “4.3.1.6 Etapas”) con tiempo de respuesta medio atípicamente corto; es decir, menor a la media general menos 3 desvíos estándares. Se excluyen así 6 participantes de 2do grado: 2 del grupo tutor (2 niños), 3 del grupo aprendiz (2 niños y 1 niña), y 1 del grupo control (1 niña).

4.3.1.3 División por grupos y parejas

Dos de las clases de 4to grado eran especialmente numerosas y para mantener el número de participantes por sesión similar al de las demás clases fueron subdivididas en dos subclases cada una. Cada una de las 11 (sub)clases fue dividida en tercios, aproximadamente. Dos de ellos fueron asignados a la condición *adulto-instructor*, y el tercero a la condición *par-instructor* (grupo *aprendiz*). A su vez, les participantes de la condición *adulto-instructor* fueron subdividides en los grupos experimentales *tutor* y *control*.

Atendiendo a diferencias en la composición por género y edad de las (sub)clases, hicimos ajustes para asegurar que estos aspectos estuvieran balanceados entre grupos experimentales por cada grado (ver Tablas 4.3 y 4.4). Además, para evitar diferencias de habilidad en la división entre *tutores* y *controles,* tuvimos en cuenta el desempeño en la etapa de Evaluación (ver sección “4.3.1.6 Etapas”), empleando un algoritmo de asignación iterativo que minimiza la diferencia de precisión entre ambos grupos por género.

Para la etapa de *Enseñanza* se formaron parejas entre participantes de los grupos *tutor* y *aprendiz*. Las parejas se formaron al azar, entre participantes de una misma (sub)clase y del mismo género.

Las asignaciones por grupo y por pareja se realizaban antes de las sesiones. Sin embargo, en caso de ausencias, se empleaba un algoritmo de decisión para realizar las reasignaciones necesarias en el momento en que la ausencia fuera notificada. El uso de este algoritmo predefinido permitía minimizar los desvíos del plan original y reducir posibles sesgos de le experimentadore ante estas situaciones impredecibles.

4.3.1.4 Procedimiento

Una plataforma experimental, dividida en *secciones* (ver Figura 4.6A y sección “4.3.1.5 Plataforma experimental”), guiaba cada una de las cuatro etapas del procedimiento: *Instrucción, Evaluación, Enseñanza/Revisión, y Reevaluación* (ver Figura 4.6B y sección “4.3.1.6 Etapas”). Cada (sub)clase participó del procedimiento durante tres días consecutivos (Figura 4.6C), y en en la medida posible procuramos que todas las (sub)clases de un mismo grado participaran los mismos tres días. Las sesiones tuvieron lugar en una de las dos salas de computación de la institución (sala mayor A, y sala menor B), y en cada una de ellas participaban alumnes de una misma (sub)clase solamente.

En resumen, el primer día (1) les participantes de la condición *adulto-instructor* tenían sus etapas de *Instrucción* y de *Evaluación* en la sala A. A continuación, les participantes de esta condición eran dividides en los grupos experimentales *control* y *tutor* (ver sección “4.3.1.3 División por grupos y parejas”). El segundo día (2) el grupo *control* tenía sus etapas de *Revisión* y de *Reevaluación* en la sala B. Simultáneamente, en la sala A, el grupo *tutor* participaba de sus etapas de *Enseñanza* y de *Reevaluación*, y el grupo *aprendiz* recibía las etapas de *Instrucción*

y de *Evaluación*, en la misma sala. Finalmente, el tercer día (3) les participantes del grupo *aprendiz* tenían sus etapas de *Revisión* y de *Reevaluación*; algunas (sub)clases en la sala A y otras en la sala B, para acelerar la toma de datos.

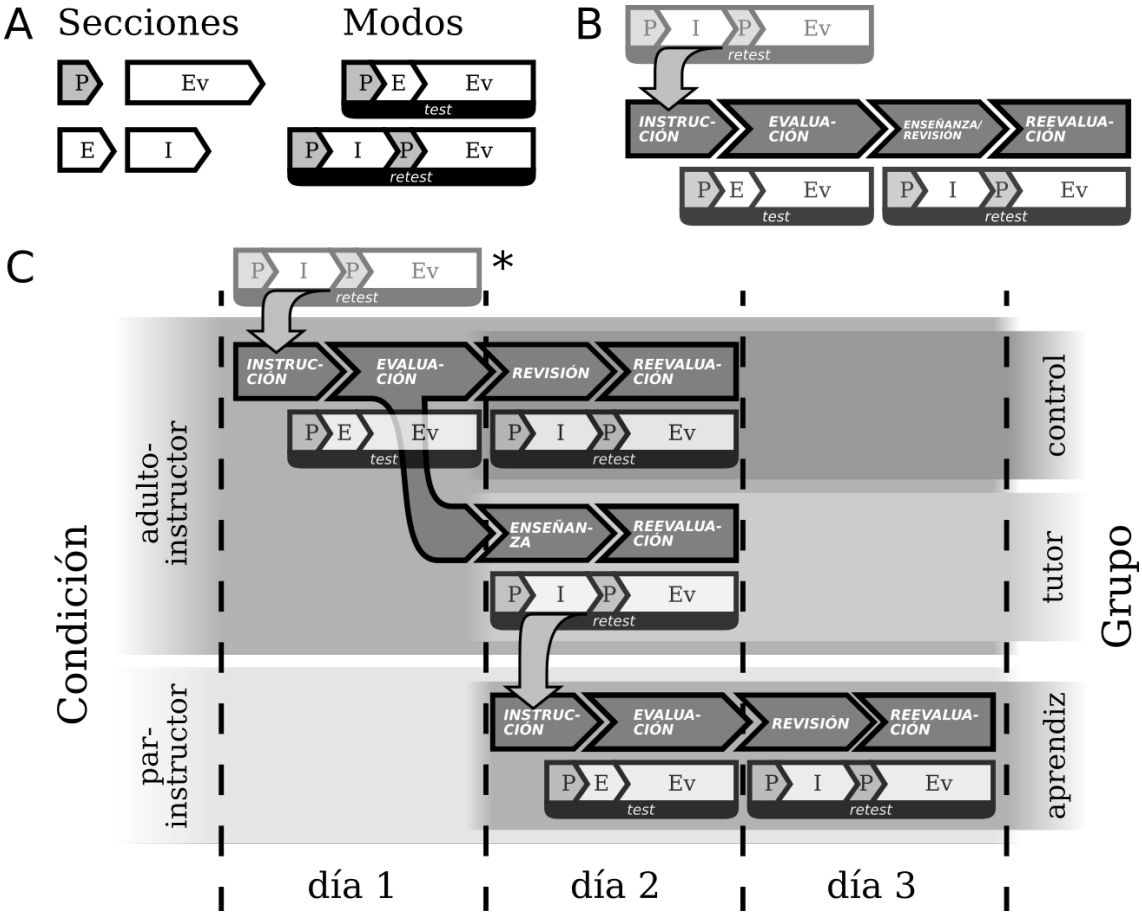


Figura 4.6: *Plataforma, etapas y procedimiento*

(A) *Plataforma experimental. Izquierda:* La plataforma experimental posee cuatro *secciones*: P: *Espera* (o *Pausa*), Ev: *Evaluación*, E: *Explicación*, e I: *Instrucción*. *Derecha*: Estas secciones se combinan y presentan secuencialmente en dos *modos* de funcionamiento de la plataforma experimental: *test* y *retest.*

(B) *Etapas*. Les participantes atendían cuatro etapas sucesivas. En la etapa de *Instrucción* aprendían las reglas de juego, enseñadas por une instructore (adulte o par) que seguía la sección *Instrucción del* modo *retest* de la plataforma experimental. Luego, en la etapa de *Evaluación*, eran evaluades por primera vez empleando el modo *test* de la plataforma. A continuación, en la etapa de *Enseñanza*/*Revisión*, utilizaban la sección *Instrucción* del modo *retest* para enseñar a une par (grupo *tutor*) o revisar a solas (grupos *control* y *aprendiz*). Finalmente, eran evaluades por segunda vez, usando la sección *Evaluación* en el modo *retest* de la plataforma.

(C) *Procedimiento.*

Día 1: Primero, en la etapa de *Instrucción* de les participantes en la condición *adulto-instructor*, une adulte experimentadore (*) usa la sección *Instrucción* del modo *retest* de la plataforma experimental para dar una clase grupal. A continuación, en la etapa de *Evaluación* de estes participantes, reciben un video explicativo y luego son evaluades usando la plataforma en modo *test*.

Día 2: Les participantes en la condición *adulto-instructor* son divides en dos grupos. El grupo *control* revisa individualmente en su etapa de *Revisión*, mientras que el grupo *tutor* enseña en su etapa de *Enseñanza*. Ésta constituye además la etapa de *Instrucción* para les participantes en la condición *par-instructor (*grupo *aprendiz*). Nuevamente, se usa la sección *Instrucción* del modo *retest* de la plataforma. A continuación, controles y tutores participan de su etapa final de *Reevaluación*, y aprendices de su etapa de *Evaluación*.

Día 3: Finalmente, emulando el grupo *control*, les participantes del grupo *aprendiz* participan de sus etapas de *Revisión* y *Evaluación*.

Las computadoras eran preparadas antes de cada sesión, procurando maximizar el espacio libre entre participantes. Por un lado, se colocaban etiquetas en los teclados numéricos para indicar la función de cada tecla (Figura 4.7A). Por otra parte, se cargaba la plataforma experimental en cada computadora, se ingresaba el código de identificación de le participante que se ubicaría en esa posición, se seleccionaba el modo de funcionamiento correspondiente, y se dejaba la plataforma retenida en la sección *Espera*, con el código de identificación en pantalla para facilitar que cada participante supiera dónde ubicarse (ver sección “4.3.1.5 Plataforma experimental”). Además, se conectaban auriculares que les participantes debían usar para minimizar las distracciones del entorno. Finalmente, antes de cada sesión, les participantes eran buscades de sus aulas y cada une recibía una etiqueta con su número de identificación, que debía adherir en su ropa.

4.3.1.5 Plataforma experimental

El *software* experimental, que desarrollamos para esta tesis, cuenta con dos *modos* de funcionamiento: “test” y “retest”, empleados en distintas etapas del procedimiento (ver sección “4.3.1.6 Etapas”). A su vez, cada *modo* de funcionamiento incluye una

serie de *secciones* (descriptas a continuación), presentadas secuencialmente. Concretamente, el modo “test” se usa en la etapa de *Evaluación,* e incluye las secciones (1) *Espera*, (2) *Explicación* y (3) *Evaluación*. Por otra parte, el modo “retest” se usa en las etapas de *Enseñanza/Revisión* y de *Reevaluación* e incluye las secciones (1) *Espera, (2) Instrucción, (3) Espera* (nuevamente) *y (4) Evaluación* (ver Figura 4.6A y Figura 4.6B).

Sección Espera

Esta sección se incluye al comienzo en ambos modos, “test” y “retest”, después de haber ingresado el código de participante y de haber seleccionado el modo de funcionamiento. Consiste de una pantalla gris, con el código de identificación de le participante bien grande para facilitar que les participantes y asistentes sepan dónde (en qué computadora) debe ubicarse cada une. Para pasar a la sección siguiente, debe introducirse un código secreto, desconocido por les participantes. En el modo “retest”, además, se incluye también en tercer lugar, entre las secciones (2) *Instrucción* y (4) *Evaluación*. Esta sección no tiene tiempo mínimo ni máximo de permanencia.

Sección Explicación

Ésta es la primera sección del modo de funcionamiento “test”, después de la sección *Espera*. En primer lugar, se presenta un audio y su transcripción correspondiente preguntando a les participantes si ya conocen el juego. Éstes deben responder “sí” o “no” usando el mouse o el teclado. Luego, se muestra un corto video explicativo con audio. En él se indica qué deberán hacer a continuación en la sección *Evaluación*, usando como ejemplo un ejercicio hipotético Set 0-diferente (Figura 4.7). Cabe aclarar que ésta es una explicación de procedimiento, en la que no se hace mención de las reglas del juego. Una vez finalizado el video, les participantes pueden seguir a la sección siguiente, o ver el video nuevamente hasta una vez más.

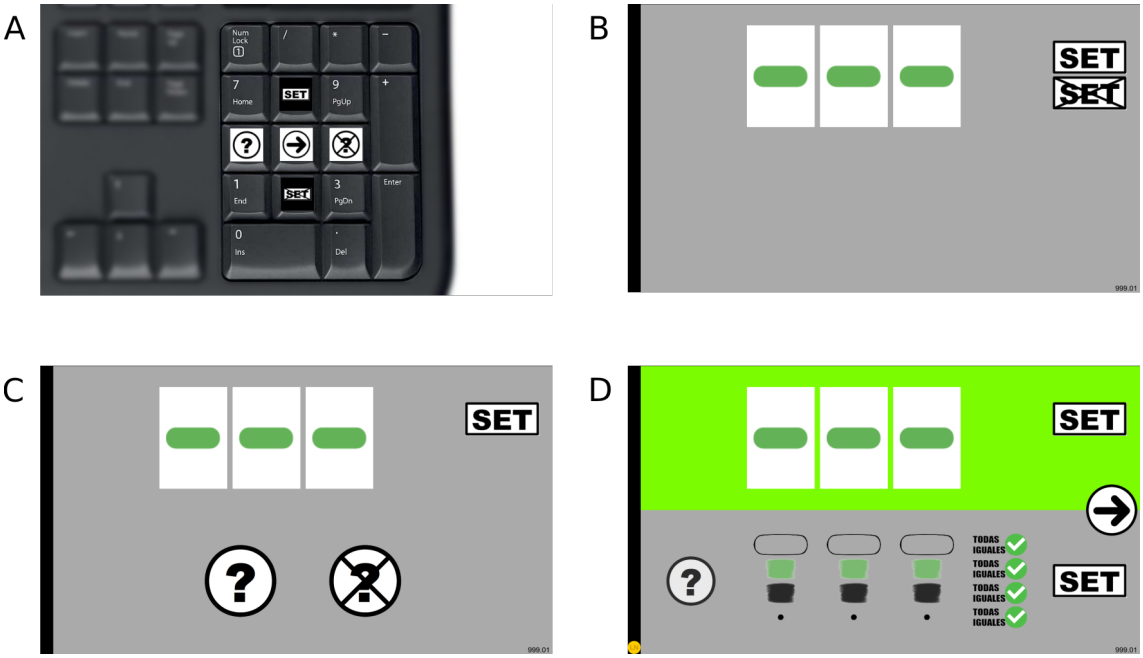


Figura 4.7: *Video explicativo*

Este video es parte de la sección *Explicación*, que se incluye en el modo de funcionamiento “test” de la plataforma experimental, utilizado en la etapa *Evaluación*. Explica qué deben hacer en los ejercicios, sin explicar las reglas del juego. Incluye cuatro pantallas, una mano que señala las partes relevantes, y el audio transcripto a continuación (entre corchetes se indican aclaraciones no verbales):

(A) *Vas a jugar al Set. Vas a jugar con el teclado.*

(B) Van a aparecer tres cartas [se señalan] y vos –con el teclado– tenés que responder si forman un Set o no [señalando las opciones en la esquina superior derecha].

(C) Tu respuesta va a aparecer acá [se señala la respuesta elegida]. Si no estás muy seguro de tu respuesta, con el teclado podés pedir una explicación [señalando la opción abajo a la izquierda]. Y si estás seguro, podés seguir sin explicación [señalando la opción abajo a la derecha]; también con el teclado.

(D) Si respondiste correctamente, la pantalla se va a pintar de verde [señalando fondo verde] y vas a ganar un punto [señalando punto ganado abajo a la izquierda]. Si no, se va a pintar de rojo. El objetivo es ganar todos los puntos que puedas. Además, si pediste explicación, acá va a aparecer la respuesta correcta [señalándola abajo a la derecha], y acá la explicación de la respuesta correcta [señalándola]. Cuando estés listo, apretá “seguir” con el teclado [señalándolo].

Sección Instrucción

Esta es la primera sección del modo de funcionamiento “retest”, después de la sección *Espera*. Comprende una guía de instrucción de cinco pantallas (ver Figura 4.8) de libre navegación; es decir que les usuaries son libres de avanzar y retroceder, y el tiempo invertido en cada pantalla queda registrado. Cabe destacar que esta sección del modo de funcionamiento “retest” es la misma que emplea le adulte instructore en la etapa de *Instrucción* (ver sección “4.3.1.6 Etapas”).

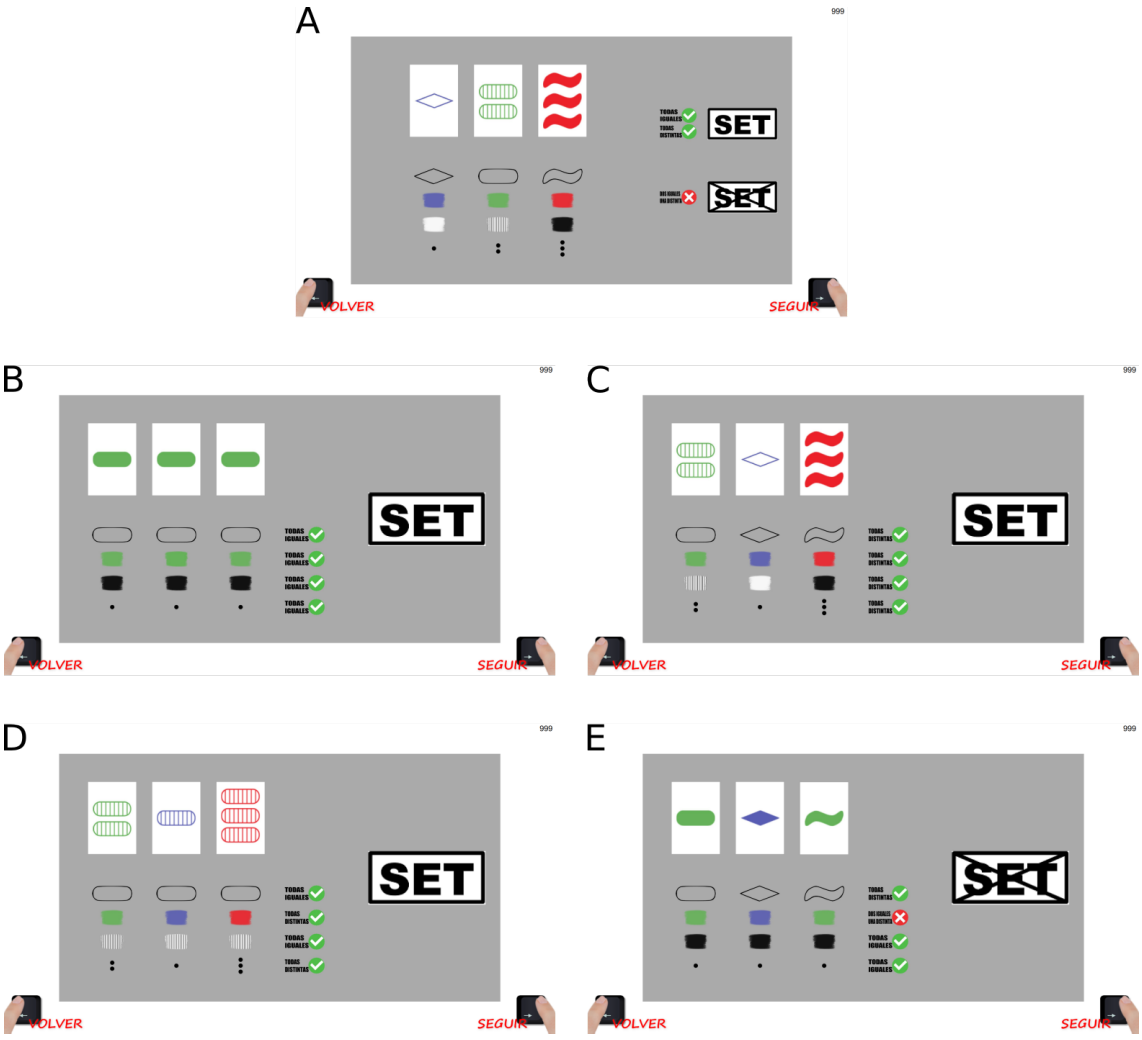


Figura 4.8: *Guía de instrucción*

Pantallas de la guía de instrucción en la sección *Instrucción* incluida en el modo “retest” de la plataforma experimental. Comprende cinco pantallas de libre navegación, cada una dividida en dos lados, tal que el lado derecho aparece a continuación del izquierdo (aquí se muestran ambos lados). Hay una sexta pantalla

que no se muestra aquí, que marca el fin de la guía de instrucción antes de seguir a la sección siguiente (de *Espera*) de la plataforma experimental.

Esta guía es utilizada por les participantes de los grupos *tutor* y *control/aprendiz*, en las etapas de *Enseñanza* y *Revisión*, respectivamente. También es usada por le adulte en la etapa de *Instrucción* de les participantes en la condición *adulto-instructor*, siguiendo un guión de instrucción (ver “Anexo 4.1: Guión de instrucción”), como se resume a continuación:

(A) En esta pantalla se presenta el juego. Se describen las cartas, los atributos y sus valores posibles. Se presenta el objetivo del juego: descubrir si las cartas forman un Set o no, y se explica la regla Set (ver sección “4.1 Introducción”, del presente capítulo).

(B) Ejemplo 1, Set 0-diferente (hipotético). Se evalúa la regla Set un atributo a la vez y se arriba a la decisión final: Set.

(C) Ejemplo 2. ídem B, con ejemplo Set 4-diferente.

(D) Ejemplo 3. Ídem B, con ejemplo Set 2-diferente.

(E) Ejemplo 4, NoSet 1-roto. Se evalúa la regla Set un atributo a la vez, hasta encontrar el atributo roto, y se arriba a la decisión final: NoSet.

Tiene un tiempo mínimo de permanencia de 100 segundos. Una vez superado este tiempo mínimo, les participantes pueden dar por concluida la explicación/revisión y avanzar a la sección siguiente desde la última pantalla. Este tiempo mínimo fue dispuesto para forzar a les participantes a considerar esta guía de instrucción, en especial en la etapa de *Revisión* (ver sección “4.3.1.6 Etapas”). Esta sección tiene además un tiempo máximo de permanencia de 10 minutos. Este tiempo es lo suficientemente largo como para ser alcanzado sólo en casos excepcionales y fue dispuesto para tener un criterio fijo y externo de cuándo la sección debe darse por concluida.

Además, en esta sección la plataforma registra el sonido ambiente. Por un lado, esto sirve para asegurar que la instrucción proporcionada por le adulte se ajusta al guión predefinido. Por otra parte, sirve también para registrar la interacción entre pares para su análisis eventual.

Sección Evaluación

Ésta es la última sección en ambos modos de funcionamiento, “test” y “retest”. En resumen, se muestra una serie de 22 ejercicios. Como se explica en detalle en el video de la sección *Explicación* (Figura 4.7), en cada ejercicio les participantes primero deben responder si las cartas exhibidas forman un Set o no. A continuación y antes de saber si respondieron correctamente, deben decidir si desean recibir una explicación de la respuesta correcta. Esta opción, además de ser una medida de la confianza, fue incluida a fin de ofrecer instancias de revisión luego de la enseñanza entre pares, visto que la ausencia de estas instancias fue identificada como posible causa de la no verificación de un efecto tutor en el Capítulo 3.

Finalmente, reciben una devolución (*feedback*) visual y auditiva de su respuesta, y la explicación de la respuesta correcta si la hubieran solicitado. La devolución auditiva busca promover que les participantes permanezcan con los auriculares colocados, para minimizar las distracciones del espacio a su alrededor. Finalmente, les participantes deciden cuándo pasar al ejercicio siguiente.

Por cada ejercicio correcto, les participantes reciben un punto. No se les dice a les participantes que hay un número limitado de ejercicios, pero se les indica que deben ganar todos los puntos que puedan. Al finalizar el último ejercicio, se les felicita por haber ganado y se les dice cuántos puntos obtuvieron. Esta sección tiene un tiempo de permanencia máximo de 10 minutos, suficientemente largo como para ser alcanzado en casos excepcionales únicamente.

Distribución de ejercicios

Los primeros dos ejercicios son de práctica y no son considerados en el análisis: un Set (hipotético) 0-diferente, y un NoSet 1-roto (color roto); siempre en el mismo orden. Ver sección “4.2.1.2 Clasificación de ejercicios” para información sobre las clases, subclases y tipos de ejercicios.

Los siguientes 20 ejercicios (de evaluación) se dividen en 4 bloques de cinco ejercicios cada uno. Cada bloque comprende un ejercicio de cada una de las subclases siguientes: Set 1-diferente, Set 3-diferente, Set 4-diferente, NoSet 3-roto, y NoSet 1-roto (ver Tablas 4.1 y 4.5). Dentro de cada bloque, el orden por subclase es aleatorio, pero una vez establecido se repite exactamente en la reevaluación para le misme participante.

En todos los casos, el grupo específico de tres cartas exhibido para cada clase, subclase y tipo de ejercicio es al azar. Cabe aclarar que no hay divisiones entre bloques para les participantes; es decir, desde su perspectiva, son 22 ejercicios consecutivos.

bloque		**Set**				**NoSet**	
		0-diferente	**1-diferente**	**3-diferente**	**4-diferente**	**3-roto**	**1-roto**
práctica	**0**	tipo único	-	-	-	-	color
prueba	**1**	-	número	textura	tipo único	textura	número
	2	-	color	forma	tipo único	forma	color
	3	-	forma	color	tipo único	color	forma
	4	-	textura	número	tipo único	número	textura

Tabla 4.5: *Distribución de ejercicios*

En los ejercicios Set 1- y 3-diferente y NoSet 1- y 3-roto, se indica el atributo extraño que define el tipo específico (dentro de cada clase y subclase). Para el Set (hipotético) 0-diferente y el Set 4-diferente hay un solo tipo posible.

Excepto en el bloque de práctica (en el que el orden de presentación es siempre primero Set 0-diferente y luego NoSet 1-roto), el orden de presentación dentro de cada bloque es aleatorio, pero una vez establecido se repite de forma idéntica en la etapa de reevaluación para le misme participante.

Para minimizar las diferencias de dificultad por bloque entre participantes (es decir, evitar que algunes participantes tuvieran, por ejemplo, un primer bloque más fácil que otres participantes), el atributo extraño para ejercicios 1-/3- diferente/roto fue predefinido para cada uno de los bloques (ver Tabla 4.5). Como se describe en detalle en el *Estudio inicial*, es posible suponer que para facilitar la identificación de Sets/NoSets es preferible reservar atributos de alta saliencia para el atributo con patrón *diferente* en Sets 1-diferente y para el atributo con patrón *roto* en NoSets 1-roto, y reservar atributos de baja saliencia para el atributo con patrón *igual* en Sets 3-diferente, y para el atributo válido (no *roto*) en NoSets 3-roto. Así, basados en los resultados obtenidos en dicho estudio, suponiendo un orden de saliencia aproximado de número > color > forma > textura (ver sección “4.2.3 Discusión” del *Estudio inicial*), e intentando que la dificultad por bloque fuese en aumento, definimos la distribución de atributo extraño por bloque y subclase de ejercicio que se muestra en la Tabla 4.5.

Aspectos técnicos

Para este estudio tradujimos el código del software desarrollado en MATLAB para el *Estudio inicial* al lenguaje de programación web JavaScript, con algunas adaptaciones específicas adicionales. Además, configuramos un servidor web en una computadora portátil del laboratorio, conectada a la red de la escuela durante las sesiones experimentales. De este modo, era posible acceder a la plataforma experimental desde las múltiples computadoras de la institución, usando un navegador ordinario, sin necesidad de instalar o configurar software adicional, maximizando así el número de participantes evaluades simultáneamente. Una vez finalizadas las sesiones, los datos eran subidos al servidor de forma automática, minimizando de esta forma el tiempo y esfuerzo necesarios para la recuperación y el almacenamiento de los datos, y evitando errores.

La plataforma de software fue diseñada con el objetivo de minimizar las posibilidades de interacción por fuera de las respuestas previstas, dentro de las limitaciones de seguridad razonables impuestas por los navegadores web. Por ejemplo, no es posible cerrar el juego sin confirmación previa (e incluso en ese caso los datos son subidos al servidor), el puntero del *mouse* es invizibilizado para evitar que les participantes intenten responder con ese dispositivo, etc. Además, todas las pantallas presentan el código de identificación de le participante y el número de ejercicio (si corresponde) en la esquina inferior derecha, para facilitar el registro de observaciones por parte de les asistentes de investigación.

El código fuente de la plataforma experimental es abierto y está disponible públicamente.[29]

4.3.1.6 Etapas

Guiades por una plataforma experimental, dividida en *secciones* (ver Figura 4.6A y sección “4.3.1.5 Plataforma experimental”), cada participante atendía cuatro etapas, descriptas a continuación (ver Figura 4.6B).

29 https://gitlab.com/diegodlh/jspsych-set

Instrucción

Les participantes en la condición *adulto-instructor* recibían una clase grupal acerca de las reglas del juego. La clase era dada por une adulte[30], siguiendo un guión (ver "Anexo 4.1: Guión de instrucción") y las pantallas de la guía de instrucción (ver Figura 4.8) provista por la sección *Instrucción* de la plataforma experimental (ver sección "4.3.1.5 Plataforma experimental"). Estas pantallas son las mismas que luego usarían les niñes para enseñar a sus compañeres o revisar, en las etapas de *Enseñanza* y *Revisión*, respectivamente. Las clases grupales tenían un tamaño de entre 4 y 11 participantes.

Por su parte, les participantes de la condición *par-instructor* eran enseñades por une par, como se describe en la etapa *Enseñanza,* más adelante.

Evaluación

Una vez que todes les participantes eran ubicades en sus respectivas posiciones asignadas, se les pedía que se colocaran los auriculares y a continuación les asistentes desbloqueaban la sección *Espera* de la plataforma experimental para dar inicio a la sección *Explicación*, en la que se mostraba un video explicando qué debían hacer a continuación. Al finalizar, comenzaba la *sección Evaluación* propiamente dicha, en la que les participantes recibían una serie de ejercicios y debían responder si las cartas exhibidas formaban un Set o no. A medida que iban finalizando, les participantes eran enviados nuevamente a sus clases.

Durante esta etapa de *Evaluación*, la forma de responder a preguntas que pudieran hacer les participantes estaba estrictamente controlada y pactada. Brevemente, les asistentes (ajenes a los detalles del estudio) respondían repitiendo la explicación del video o parte de ella, solamente, señalando las partes de la pantalla y/o del teclado que correspondieran.

Los resultados de esta etapa se usaron para dividir a les participantes en la condición *adulto-instructor* en los grupos experimentales *tutor* y *control*, tal que ambos grupos estuvieran balanceados en género, edad y desempeño (ver sección "4.3.1.3 División por grupos y parejas").

30 Todas las clases eran dadas por le misme adulte, de género masculino.

Enseñanza/Revisión

Por un lado, les participantes del grupo *tutor* participaban de la versión *Enseñanza* de esta etapa. Cada participante ocupaba una computadora, junto a su pareja preasignada del grupo *aprendiz* (ver sección “4.3.1.3 División por grupos y parejas”), con le tutore a la derecha, cerca del teclado numérico, y le aprendiz a la izquierda. A continuación, en voz alta se felicitaba a les tutores por su participación el día anterior y se les indicaba que esta vez debían enseñar a sus compañeres así elles podían jugar después también. A continuación, se iniciaba la *sección Instrucción* de la plataforma en cada una de las computadoras; en esta sección le tutore podía avanzar y retroceder libremente a través de las mismo cinco pantallas de la guía de instrucción (ver Figura 4.8) utilizadas por le instructore adulte en la etapa de Instrucción. A medida que les tutores manifestaban haber concluido la explicación, le aprendiz correspondiente era conducide a otra computadora en la mitad opuesta de la sala para su etapa de *Evaluación*. Por otra parte, le tutor permanecía en la misma computadora y se le solicitaba que se coloque los auriculares para su etapa de *Reevaluación*.

Por otra parte, les participantes del grupo *control*, y les del grupo *aprendiz* al día siguiente (ver sección “4.3.1.4 Procedimiento”*)*, recibían la versión *Revisión* de esta etapa. En voz alta, eran felicitades por su participación el día anterior y se les indicaba que debían revisar las instrucciones individualmente para jugar todavía mejor la próxima vez. A continuación, se les solicitaba que se colocaran los auriculares y la sección *Instrucción* de la plataforma era iniciada en cada computadora. A medida que les participantes manifestaban haber concluido con la revisión, permanecían en la misma ubicación para la etapa de *Reevaluación*.

En esta etapa les asistentes controlaban que les participantes (o parejas) no interactuaran entre elles, pero manteniendo una distancia prudente para no hacerles sentirse observades. En caso de consultas, se limitaban a explicar cómo navegar a través de las pantallas de la guía de instrucción, dirigiéndose específicamente a le tutore en la versión *Enseñanza*.

Reevaluación

Esta etapa es idéntica a la de *Evaluación*, pero sin el video explicativo al comienzo. Empleaba la sección *Evaluación* del modo “retest” de la plataforma experimental,

con el mismo orden de presentación de los ejercicios que el recibido en la etapa de *Evaluación*, pero con ejercicios isomorfos; es decir, de la misma clase, subclase y tipo, pero diferentes en cuanto a la combinación específica de valores para cada atributo (ver sección “4.3.1.5 Plataforma experimental”). Al igual que en la etapa de *Evaluación*, a medida que iban finalizando les participantes eran enviades a sus clases nuevamente.

4.3.1.7 Metacognición

Al final de la etapa de *Reevaluación*, algunes participantes recibieron una prueba de metacognición, para evaluar la relación entre esta habilidad y la enseñanza entre pares. El procedimiento empleado es una adaptación del utilizado previamente en nuestro laboratorio (Salles et al., 2016), que a su vez fue adaptado del publicado por Kornell y colaboradores (2007).

Procedimiento

Brevemente, cada ejercicio de la prueba de metacognición se divide en una tarea de tipo I, en la que les participantes deben estimar cuál de 9 círculos presentados en pantalla es el más grande; y una tarea de tipo II, en la que deben reportar si creen haber respondido correctamente en la tarea de tipo I (alta confianza) o no (baja confianza). En cada ejercicio les participantes reciben puntos en función de sus respuestas. Si reportan alta confianza, reciben 3 puntos si habían respondido correctamente, pero pierden 3 si no. En cambio, si reportan baja confianza, reciben siempre 1 punto, sin importar su respuesta en la tarea de tipo I. Por otra parte, para evaluar la habilidad metacognitiva independientemente del desempeño en la tarea de tipo I, la dificultad de esta tarea se va ajustando automáticamente según el desempeño de cada participante, de modo que todes tengan una precisión del 75%, independientemente de su habilidad.

A diferencia del protocolo original, en que une experimentadore guiaba individualmente a cada participante, en este estudio evaluamos múltiples participantes simultáneamente, que respondían a solas usando el *mouse* de la computadora asignada. Por este motivo, desarrollamos una adaptación en JavaScript para incluirla a continuación de la etapa de *Reevaluación*,[31] usando el

31 La adaptación en JavaScript de la plataforma experimental de la prueba de metacognición fue desarrollada por Ariel Haimovici.

mismo navegador web y aprovechando las mismas ventajas de guardado automático en el servidor descriptas en el apartado *Aspectos técnicos* de la sección "4.3.1.5 Plataforma experimental". Además, como les participantes concluían la etapa –anterior– de *Reevaluación* a distintos tiempos, no podíamos explicar la tarea verbalmente. En cambio, les participantes recibieron instrucciones escritas en pantalla. Además, atendiendo al corto tiempo del que disponíamos, debimos omitir los ejercicios de ejemplo incluidos en la explicación original. El número de ejercicios era variable en ambos casos y el promedio fue similar (68±20 en este estudio, 60±7 en el original).

Por razones logísticas (concretamente, demoras en la preparación de la prueba), les participantes de 2do grado y les de una de las subclases de 4to (4C1) no pudieron ser evaluades.

Análisis

Para analizar los datos obtenidos aplicamos un modelo de Teoría de Detección de Señales a los datos obtenidos en la prueba de metacognición,[32] como se describe en detalle en una publicación previa de nuestro laboratorio (Salles et al., 2016). Este modelo clasifica las respuestas de les participantes en cuatro categorías en una tabla de contingencia, según hayan respondido correcta o incorrectamente, y reportado alta o baja confianza. De este modo es posible discriminar y obtener medidas de dos aspectos del juicio metacognitivo: (1) la habilidad metacognitiva propiamente dicha, y (2) la aversión al riesgo. En particular, la habilidad metacognitiva, es decir la capacidad de diferenciar respuestas correctas e incorrectas en la tarea de tipo I, se mide con el parámetro *d'* del modelo, que expresa la separación entre dos hipotéticas distribuciones normales, correspondientes a haber respondido correcta o incorrectamente la tarea de tipo I, respectivamente, y de las que le sujeto muestrearía una señal metacognitiva interna que le llevaría a reportar alta o baja confianza en la tarea de tipo II (Fleming & Lau, 2014; Salles et al., 2016).

Como mencionamos arriba, para obtener un valor del parámetro *d'* independiente del desempeño en la tarea de tipo I, es fundamental controlar el desempeño en esta

32 El análisis de los datos de metacognición fue realizado por Julieta Goldstein, bajo la dirección conjunta del tesista, Diego de la Hera, y su codirectora, Cecilia Calero.

tarea, ajustándola automáticamente al 75% de precisión. Sin embargo, debido a un error en el código de nuestra adaptación, para algunes participantes el desempeño no logró estabilizarse en 75%. Además, en nuestro estudio la estabilidad se alcanzó en promedio un poco más tarde (ejercicio 5) que en el estudio original (ejercicio 3), y como la tarea era autoadministrada es posible que les participantes hayan perdido interés más rápidamente y que los últimos ejercicios de algunas sesiones no sean representativos. Por tanto, para decidir con qué datos trabajar, para cada ejercicio se calculó el desempeño promedio de la ventana de 16 ejercicios que incluye dicho ejercicio, los 8 anteriores y los 7 siguientes. Se conservaron únicamente los ejercicios cuya ventana tiene un desempeño promedio entre 62,5 y 87,5% (75% ± 12,5%). Participantes con más del 50% de ejercicios descartados no fueron considerados para el análisis: 3 de 4to grado (2 niñas y 1 niño), 8 de 6to grado (5 niñas y 3 niños). De les participantes restantes, en promedio 76% de los ejercicios fueron conservados. El resto del análisis fue realizado según los trabajos previos de nuestro laboratorio (Salles et al., 2016), incluyendo la estrategia de asignar un “1” en la celda correspondiente de la tabla de contingencia en sesiones sin ejercicios para alguna de las cuatro categorías.

Los datos de una participante de 4to grado del grupo *aprendiz,* que no asistió a clase el día de su etapa de *Reevaluación*, y los de dos participantes (una de 4to y uno de 6to), que no se guardaron por un error técnico, no pudieron ser incluidos en el análisis.

4.3.1.8 Análisis

En todos los casos en los que se realizaron pruebas paramétricas, se verificó el cumplimiento de todos los supuestos correspondientes, incluyendo la distribución normal de los datos (con prueba de Shapiro-Wilk), y la homogeneidad de varianzas (con prueba de Bartlett). En los casos en los que estos supuestos no se cumplieran, se usaron test no paramétricos alternativos, indicándose esto en el texto.

El análisis estadístico de los datos y las visualizaciones se llevaron a cabo con Pandas, Seaborn, StatsModels, SciPy, afex, y otras librerías de Python y R, en Jupyter Notebook.[33]

33 https://pandas.pydata.org/ • https://seaborn.pydata.org/ • https://www.statsmodels.org/ • https://www.scipy.org • https://afex.singmann.science/ • https://jupyter.org/.

4.3.2 Resultados

4.3.2.1 Efecto del tipo de instrucción

En primer lugar comparamos a les participantes en función del tipo de instrucción recibida. Es decir, si fueron enseñades por une adulte (condición *adulto-instructor*) o por une par (condición *par-instructor*). Por un lado, evaluamos dos indicadores de desempeño: (1) la precisión, es decir la proporción de respuestas correctas, y (2) el tiempo de respuesta; por otra parte, evaluamos (3) la confianza.

Tipo de instrucción y precisión

Para este análisis, usamos los datos de la etapa de *Evaluación* solamente. En primer lugar, realizamos un ANOVA de 2 vías, con precisión media como variable dependiente, y grado y condición (*adulto-instructor* y *par-instructor*) como factores. Se observa un efecto significativo de grado, $F(2, 112) = 13{,}7$, $p < 1{,}00\text{E-}4$, y de condición, $F(1, 112) = 5{,}01$, $p = 0{,}03$, pero sin efecto de interacción grado:condición, $F(2, 112) = 1{,}43$, $p = 0{,}24$ (ver Figura 4.9A). Un análisis post-hoc por grado con prueba *HSD* de Tukey muestra que la diferencia de precisión por condición (a favor de la condición *adulto-instructor*) es significativa en 4to grado, $p = 8{,}20\text{E-}3$, y no significativa en 2do, $p = 0{,}935$, y 6to grados, $p = 9{,}62\text{E-}2$. Cabe destacar que la precisión en ambas condiciones para todos los grados resulto significativamente mayor que la esperada por azar (50% de precisión), como surge del análisis post-hoc con corrección de Bonferroni por comparaciones múltiples.

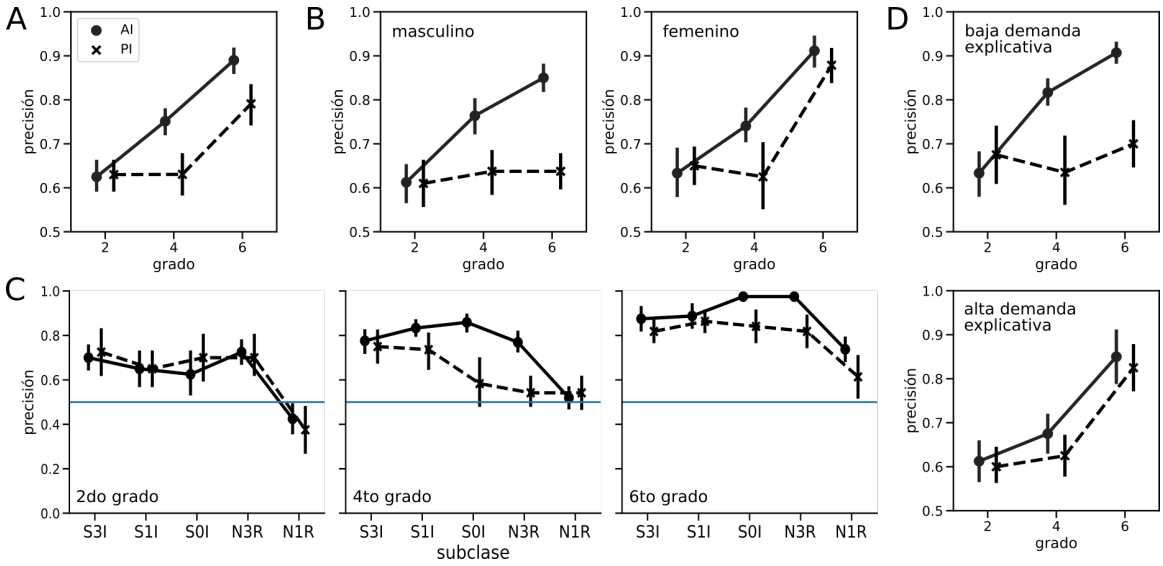


Figura 4.9: *Precisión*

–●–: adulto-instructor (AI); –×–: par-instructor (PI)

(A) Precisión por grado, en función del tipo de instrucción. Les participantes enseñades por une adulte fueron más precises que les enseñades por une par. Esta diferencia es significativa en 4to grado. Además, la precisión aumenta con el grado, pero esta dependencia es igual para ambas condiciones.

(B) División por género. La diferencia entre condiciones experimentales, observada en la Figura A, se verifica sólo en participantes de género masculino (panel izquierdo), para 4to y 6to grado. Por el contrario, no se observan diferencias de precisión entre aprender de une adulte y aprender de une par en participantes de género femenino (panel derecho).

(C) División por subclase de ejercicio. Las diferencias de precisión en 4to grado entre participantes enseñades por une adulte o por une par (Figura A), se deben a ejercicios de dificultad intermedia de las subclases Set 4-diferente y NoSet 3-roto. SXI: “Set X-igual”, NXR: “NoSet X-roto”.

(D) División por demanda de explicación. Antes de saber si habían respondido correctamente, les participantes podían solicitar una explicación del ejercicio. Cuando (dentro de cada grado y género) se consideran solamente les participantes que piden pocas explicaciones (panel superior), la diferencia de precisión entre condiciones observada se mantiene, en 4to y 6to grado. Por el contrario, cuando se consideran les participantes que piden explicación más frecuentemente (panel inferior), esta diferencia desaparece.

A continuación repetimos el mismo análisis dividiendo los datos según (1) el género de les participantes, (2) la subclase de ejercicio, y (3) la demanda de explicación.

División por género

En el caso de los participantes de género masculino, se observa nuevamente un efecto significativo de condición (*adulto-instructor* vs *par-instructor*), $F(1, 44) = 6{,}71$, $p = 0{,}01$, y una prueba post-hoc *HSD* de Tukey muestra que la diferencia de precisión entre condiciones es significativa en 4to, $p = 3{,}80\text{E-}2$, y 6to grado, $p = 1{,}88\text{E-}2$. Sin embargo, en el caso de las participantes de género femenino, no se observa diferencia significativa entre condiciones, $F(1, 62) = 0{,}95$, $p = 0{,}33$, y sólo se observa efecto significativo de grado, $F(2, 62) = 11{,}8$, $p < 1{,}00\text{E-}4$. Es decir, la diferencia de precisión entre participantes enseñades por une adulte y participantes

enseñades por une par, reportada en el párrafo anterior, no se observa en participantes de género femenino (Figura 4.9B).

División por subclase de ejercicio

Luego, analizamos si la brecha de precisión, observada entre participantes enseñades por une adulte y otres enseñades por une par, es generalizada o se debe a subclases de ejercicio específicas (ver Tabla 4.1). Para esto, se realizaron múltiples pruebas no paramétricas de Mann-Whitney, una por cada subclase de ejercicio, con valores *p* ajustados por comparaciones múltiples (cinco comparaciones por grado) según el método de Holm-Sidak (ver Tabla 4.6). Los resultados de estos análisis muestran que les participantes enseñades por une adulte son significativamente más precisos que les enseñades por une par solamente en 4to grado y solamente en ejercicios de dificultad intermedia de las subclases Set 4-diferente, $p = 0{,}031$, y NoSet 3-roto, $p = 0{,}011$. Es posible que en ejercicios más fáciles (Set 1-diferente y 3-diferente) y más difíciles (NoSet 1-roto) la precisión fuese demasiado alta o baja, respectivamente, para notar diferencias (ver Figura 4.9C). La estimación de dificultad relativa de los ejercicios surge de los resultados del *Estudio inicial*.

Clase	Subclase	p valor ajustado		
		2do grado	4to grado	6to grado
Set	1-diferente	0,814	0,492	0,292
	3-diferente	0,814	0,468	0,292
	4-diferente	0,814	0,031	0,292
NoSet	3-roto	0,814	0,011	0,054
	1-roto	0,814	0,492	0,292

Tabla 4.6: *Diferencia de precisión entre condiciones por subclase de ejercicio*

Se comparan las condiciones *adulto-instructor* vs *par-instructor* en la etapa de Evaluación. p valores de pruebas no paramétricas de Mann-Whitney, una por cada subclase de ejercicio, con valores p ajustados por comparaciones múltiples (cinco comparaciones por grado) según el método de Holm-Sidak.

División por demanda de explicación

Luego de responder cada ejercicio y antes de saber si lo habían hecho correctamente, les participantes podían solicitar una explicación de la respuesta correcta (ver apartado *Sección Evaluación* de la sección “4.3.1.5 Plataforma

experimental"). ¿Es posible que una mayor tendencia a pedir explicación pueda haber cerrado la brecha de precisión entre participantes enseñades por une adulte y aquelles enseñades por une par? Para responder esta pregunta, dividimos a les participantes en bajes y altes solicitantes de explicación, según la mediana por grado y género, y repetimos el análisis del primer párrafo. Para el grupo de baja solicitud de explicaciones, se observa nuevamente un efecto significativo de condición sobre la precisión, $F(1, 58) = 5{,}72$, $p = 0{,}02$, con diferencias significativas entre condiciones en 4to, $p = 3{,}33\text{E-}3$, y 6to grados, $p = 3{,}91\text{E-}2$, según una prueba post-hoc *HSD* de Tukey. Sin embargo, para el grupo de alta solicitud de explicaciones, el efecto de condición no es significativo, $F(1, 48) = 0{,}42$, $p = 0{,}52$ (ver Figura 4.9D). Si bien preliminares, en conjunto estos resultados sugieren que una mayor propensión a pedir explicación podría reducir la brecha de precisión entre les participantes enseñades por une adulte y aquelles enseñades por une par.

Tipo de instrucción y tiempo de respuesta

El otro indicador de desempeño analizado fue el tiempo de respuesta, sin discriminar entre respuestas correctas e incorrectas. Realizamos un ANOVA de 3 vías, con grado, condición y género como factores, y el logaritmo del tiempo de respuesta como variable dependiente. Se observa un efecto significativo de condición, $F(1, 106) = 8{,}40$, $p = 0{,}005$, sin efecto significativo de género, $F(1, 106) = 2{,}94$, $p = 0{,}09$, o de la interacción grado:condición, $F(2, 106) = 2{,}28$, $p = 0{,}11$. Una prueba post-hoc HSD de Tukey muestra que el tiempo de respuesta promedio para participantes enseñades por une par es significativamente menor que para aquelles enseñades por une adulte (Figura 4.10). Este resultado, similar al obtenido en el estudio principal del Capítulo 3, podría explicar en parte la menor precisión de les participantes enseñades por une par descripta en la sección anterior.

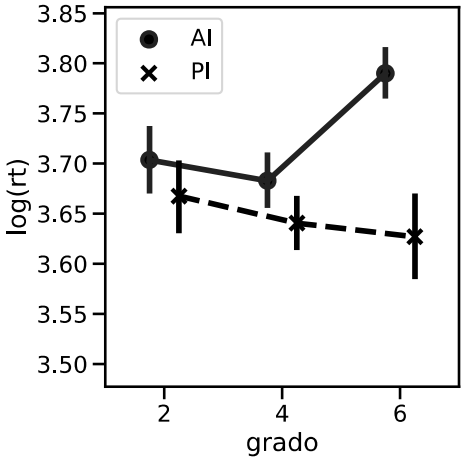


Figura 4.10: *Tiempo de respuesta*

Les participantes enseñades por une par (PI: *par-instructor*) respondieron en promedio más rápido que les enseñades por une adulte (AI: *adulto-instructor*). rt: tiempo de respuesta.

Tipo de instrucción y confianza

Además de comparar el desempeño de les participantes según el tipo de instrucción, comparamos también su confianza. Es decir, evaluamos si les participantes enseñades por une par y aquelles enseñades por une adulte diferían en cuanto a su inclinación por pedir explicación de los ejercicios (ver apartado *Sección Evaluación* de la sección “4.3.1.5 Plataforma experimental”).

Realizamos un ANOVA de tres vías, con proporción de ejercicios explicados como variable dependiente, y grado, condición y género como factores. Sin embargo, no observamos efecto significativo de condición, $F(1, 106) = 1{,}76$, $p = 0{,}19$. Por otra parte, aunque los efectos de grado, $F(2, 106) = 3{,}34$, $p = 0{,}04$, y género, $F(1, 106) = 6{,}19$, $p = 0{,}01$, resultaron significativos, no son válidos por no cumplirse los supuestos de normalidad y homogeneidad de varianzas de esta prueba paramétrica. Por tanto, a continuación realizamos una prueba no paramétrica de Kruskal-Wallis, enfocándonos en la variable género. Sin embargo, no encontramos un efecto significativo, $H(1) = 0{,}786$, $p = 0{,}375$. Resultados similares fueron obtenidos realizando pruebas independientes por condición, o por condición y género.

Por lo tanto, no parece haber diferencias en la confianza, medida como tendencia a pedir explicación, entre participantes enseñades por une adulte y sus compañeres enseñades por une par, ni diferencias significativas por grado o género tampoco.

4.3.2.2 Efecto de enseñar

Finalmente, evaluamos el efecto de enseñar, en comparación con repasar a solas. Nuevamente evaluamos por un lado (1) el desempeño, y por otra parte (2) la confianza.

Enseñar y desempeño

En esta oportunidad, para la evaluación del desempeño nos enfocamos en la precisión, en la etapa de *Reevaluación*. Realizamos un ANOVA de tres vías, con la diferencia de precisión entre evaluación y reevaluación como variable dependiente. No observamos efectos significativos de grado, $F(2, 67) = 0{,}25$, $p = 0{,}78$, grupo (*control* vs *tutor*), $F(1, 67) = 1{,}37$, $p = 0{,}25$, o género, $F(1, 67) = 0{,}09$, $p = 0{,}76$ (Figura 4.11).

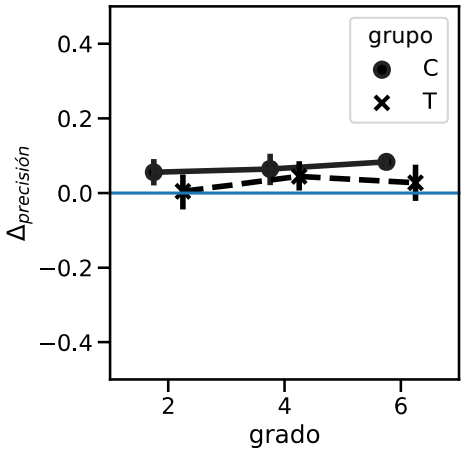


Figura 4.11: *Efecto de enseñar*

No se observa diferencia entre participantes que enseñaron a une par (T: *tutor*) y aquelles que repasaron a solas (C: *control*) en cuanto al cambio de precisión entre la evaluación y la reevaluación (Δ*precisión*).

Modulación metacognitiva

A continuación evaluamos el efecto modulador que la habilidad metacognitiva de les participantes podría tener sobre sus aprendizajes mediados por las etapas de

enseñanza y revisión; es decir, sobre las diferencias de precisión entre evaluación y reevaluación.[34]

En primer lugar comparamos la distribución de valores del parámetro *d'* (habilidad metacognitiva; ver apartado "4.3.1.7 Metacognición" de la sección *Métodos*) con los resultados obtenidos en trabajos previos (Salles et al., 2016). Si bien la habilidad metacognitiva media en este estudio, *d*' = 0,817±0,701, resultó menor que la obtenida anteriormente, *d*' = 1,09±0,479, esta diferencia no es significativa, según una prueba no paramétrica de Mann-Whitney, $U = 964$, $p = 0{,}060$.

Por otra parte, realizamos un ANOVA de tres vías, con habilidad metacognitiva (*d'*) como variable dependiente, y grado, grupo y género como factores. No observamos efecto significativo de grupo, $F(1, 51) = 1{,}61$, $p = 0{,}209$. Sin embargo, sí observamos un efecto significativo de grado, con habilidad metacognitiva en 4to, *d*' = 0,720±0,699, significativamente menor que en 6to, *d*' = 0,965±0,631, $F(1, 51) = 8{,}15$, $p = 0{,}006$; y un efecto significativo de género, con habilidad metacognitiva en niñas, *d*' = 0,583±0,521, significativamente menor que en niños, *d*' = 1,12±0,762, $F(1, 51) = 16{,}8$, $p < 0{,}001$. Comentamos las posibles razones y los alcances de estas diferencias en la sección "4.3.3 Discusión".

A continuación, para evaluar el efecto modulador de la metacognición, realizamos regresiones de modelo lineal mixto (una por grado) con diferencia de precisión entre evaluación y reevaluación ($\Delta_{precisión}$) como variable dependiente, habilidad metacognitiva (*d'*) como variable independiente, y grupo (*tutor* vs *control*+*aprendiz*) como factor. Los grupos *control* y *aprendiz* fueron considerados como un grupo único para simplificar el análisis porque ambos difieren del grupo *tutor* en haber revisado a solas en vez de haber enseñado a une compañere. La dependencia entre la diferencia de precisión ($\Delta_{precisión}$) y la habilidad metacognitiva (*d'*) no resultó significativamente diferente entre grupos, ni para 4to ni para 6to grado (ver Tabla 4.7). Por otra parte, realizamos correlaciones de Spearman (una por grado y grupo) entre diferencia de precisión y habilidad metacognitiva. Si bien el coeficiente de correlación resultó positivo para el grupo *tutor* en 4to grado, $r = 0{,}339$, ninguno de los coeficientes obtenidos resultó significativo (ver Tabla 4.7).

34 El análisis de los datos de metacognición fue realizado por Julieta Goldstein, bajo la dirección conjunta del tesista, Diego de la Hera, y su codirectora, Cecilia Calero.

Prueba	Grupo	Coef./Valor	Grado	
			4	6
RMLM (*)	CA	b	-0,008	-0,005
	T	b	0,103	-0,046
		z (**)	1,374	-0,397
		p (**)	0,169	0,691
Correlación Spearman	CA	r	0,121	-0,108
		p	0,547	0,690
	T	r	0,339	-0,273
		p	0,236	0,600

Tabla 4.7: *Modulación metacognitiva del efecto de enseñar sobre el desempeño*

* regresión de modelo lineal mixto: $\Delta_{precisión} = a + b * d'$, *con* $\Delta_{precisión}$ *la diferencia de precisión entre evaluación y reevaluación, y d' la habilidad metacognitiva.*

** valores z y p para el contraste del coeficiente b entre grupo tutor (T) y control+aprendiz (CA).

Es decir, si bien parece haber una tendencia de la habilidad metacognitiva como moduladora del efecto de enseñar sobre el desempeño en 4to grado, este resultado no es significativo.

Enseñar y confianza

Por último, además de evaluar el efecto de enseñar sobre la diferencia de desempeño, evaluamos también su efecto sobre la confianza, medida aquí como la tendencia a pedir explicación de los ejercicios (ver apartado *Sección Evaluación* de la sección "4.3.1.5 Plataforma experimental").

Realizamos un ANOVA de tres vías, con la diferencia de proporción de ejercicios explicados entre evaluación y reevaluación como variable dependiente, y grado, grupo y género como factores. No encontramos efecto significativo de grado, $F(2, 67) = 1,77$, $p = 0,18$; grupo, $F(2, 67) = 0,01$, $p = 0,92$; o género, $F(2, 67) = 0,87$, $p = 0,35$. Además, como no se cumplen los supuestos de esta prueba paramétrica, realizamos también pruebas no paramétricas de Kruskal-Wallis, una por grado, enfocándonos en el factor grupo. Sin embargo, ninguna resultó significativa, $H(1) = 9,64E\text{-}2$, $p = 0,756$; $H(1) = 1,96$, $p = 0,162$; y $H(1) = 0,634$, $p = 0,426$; para 2do, 4to y 6to grado, respectivamente.

4.3.3 Discusión

El objetivo de este estudio fue evaluar el impacto de la enseñanza entre pares sobre el aprendizaje de tutores y aprendices en una tarea que combina cuestiones cognitivo/conceptuales y percepción visual, y su interacción con variables demográficas y cognitivas de les participantes, específicamente la edad, el género y la metacognición.

En primer lugar evaluamos la efectividad de la enseñanza entre pares sobre les aprendices, comparando el desempeño en la tarea entre participantes que habían aprendido de une adulte y participantes que habían aprendido de une par. Consideramos dos medidas del desempeño: la precisión o porcentaje de respuestas correctas, y el tiempo de respuesta.

Primero observamos que si bien en promedio todes les participantes exhibieron una precisión mayor a la esperada por azar, quienes habían aprendido de une adulte fueron significativamente más precises que quienes habían aprendido de une par. Esta mayor precisión se habría dado a expensas de mayores tiempos de respuesta, revelando un posible equilibrio entre velocidad y precisión similar al discutido en el Capítulo 3. Cabe señalar que, como comentamos en la sección “4.2.3 Discusión” del *Estudio inicial*, un análisis alternativo podría haber considerado la variable *eficiencia inversa*, que combina ambas medidas. Sin embargo, su uso estaría desaconsejado en casos como éste, en que precisión y velocidad cambian en sentido opuesto entre condiciones (Townsend & Ashby, 1983).

Por otra parte, con un análisis más detallado, observamos que esta diferencia de precisión se verifica sólo en 4to y 6to grado, y sólo en ejercicios de dificultad intermedia. Es posible que en los demás casos los ejercicios hubieran resultado demasiado difíciles o demasiado fáciles como para observar diferencias por tipo de instrucción. Por otra parte, esta diferencia entre aprender de une par y aprender de une adulte se observa en participantes varones solamente.

Además, si se toma en consideración que en cada ejercicio, después de responder y antes de saber si lo habían hecho correctamente, les participantes podían solicitar una explicación de la respuesta correcta, y se divide a les participantes en altes y bajes solicitantes de explicación, observamos que la diferencia por tipo de instrucción no existe en el grupo de altes solicitantes. Si bien la diferencia entre

géneros en la tendencia a pedir explicación no es significativa, no podemos descartar que una posible mayor tendencia de las niñas a pedir explicación explique el cierre de la brecha de precisión por tipo de instrucción observada en este género, e investigaciones futuras enfocadas en esta cuestión podrían echar luz a esta posibilidad.

En segundo lugar, evaluamos el impacto de la enseñanza entre pares sobre el aprendizaje de les propies tutores (efecto tutor). Por un lado, a diferencia de los estudios descriptos en los dos capítulos anteriores, indicábamos a les participantes si habían respondido correctamente o no. Por otra parte, les participantes podían optar por recibir una explicación (a modo de recordatorio) a continuación de cada uno de los ejercicios (ver apartado *Sección Evaluación* de la sección “4.3.1.5 Plataforma experimental”). No obstante, no observamos diferencias de desempeño en la etapa de reevaluación entre quienes enseñaron a une par (grupo *tutor*) y quienes revisaron a solas (grupo *control*), para ninguna de las edades y géneros evaluados. Cabe mencionar, sin embargo, que al igual que los estudios de los Capítulos 2 y 3, tampoco observamos perjuicios para el grupo *tutor*.

Por último, nos enfocamos en la relación entre la enseñanza entre pares y la metacognición. En primer lugar empleamos autoreportes de confianza acerca del desempeño en la tarea de discriminación entre Sets y NoSets. Brevemente, en cada ejercicio, luego de responder y antes de saber si habían respondido correctamente, les participantes podían indicar que no estaban seguros (baja confianza) y que deseaban recibir una explicación de la respuesta correcta, o que estaban seguros (alta confianza) y que deseaban continuar sin explicación (ver apartado *Sección Evaluación* de la sección “4.3.1.5 Plataforma experimental”). Estos reportes son una medida indirecta de procesos metacognitivos de monitoreo de la incertidumbre. Sin embargo, en este estudio, al estar necesariamente acoplados al pedido de explicación adicional, es probable que involucren también procesos metacognitivos de control (Ghetti et al., 2013). Al igual que en el estudio del Capítulo 3, comparamos participantes que habían enseñado a une par (grupo *tutor*) y participantes que habían revisado a solas (grupo *control*). Sin embargo, en este caso no observamos que el cambio de confianza de la evaluación a la reevaluación fuese diferente entre ambas condiciones, sugiriendo que la enseñanza no habría tenido efectos metacognitivos evaluables con el paradigma empleado.

Por otra parte, la naturaleza automática y simultánea de la administración de las evaluaciones en este estudio nos permitió considerar una medida más adecuada y directa de la metacognición. De este modo pudimos analizar cómo ésta modula los aprendizajes derivados de la enseñanza entre pares, un aspecto de la relación entre ambas que no habíamos podido abordar en los estudios anteriores. Puntualmente, en 4to grado observamos una posible relación entre habilidad metacognitiva y el aprendizaje derivado de la etapa de enseñanza para el grupo *tutor*, en contraste con el derivado de la etapa de revisión para los grupos *control* y *aprendiz*. Si bien esta asociación no resultó estadísticamente significativa, ni fue observada en 6to grado, estos resultados preliminares subrayan la necesidad de realizar estudios futuros que establezcan si participantes con mayor habilidad metacognitiva podrían aprovechar mejor estas instancias de enseñanza a une par.

Con respecto al método empleado para medir metacognición, es importante tener en cuenta que los resultados obtenidos, si bien semejantes, no son idénticos a los obtenidos en trabajos anteriores del laboratorio (Salles et al., 2016). En particular, observamos diferencias por edad (grado) y género que no habíamos observado anteriormente. Cabe mencionar que el procedimiento empleado en este estudio es una versión modificada del original, atendiendo a las características particulares del estudio; particularmente el tamaño de la muestra y el tiempo disponible. Concretamente, hubo que excluir los ejercicios de ejemplo de la explicación, por falta de tiempo, y la explicación y administración de la prueba fueron autoguiadas, es decir sin la asistencia personalizada de une investigadore. Además, la prueba no era administrada de forma aislada, sino después de las etapas de enseñanza/revisión y reevaluación. Es posible que estas diferencias hayan impactado en la comprensión de la tarea, en las estrategias de resolución, y en la atención y motivación de les participantes. Es comprensible que el impacto de estas diferencias de administración sobre el resultado final dependa de factores como la edad de les participantes, explicando así parte de las diferencias observadas en los resultados.

Si, en efecto, la medida (*d'*) de metacognición está "contaminada" por factores ajenos, es importante interpretar los resultados con cautela. Además, si bien la prueba autoadministrada empleada en este estudio podría ser una alternativa adecuada a la forma guiada empleada en trabajos previos, en trabajos futuros sugerimos ajustar la explicación y administrar la prueba de forma aislada, para

minimizar el impacto que las diferencias de administración podrían tener sobre el resultado final.

4.4 Conclusiones

En los capítulos anteriores abordamos la cuestión de la enseñanza entre pares en dominios de relevancia para el mundo presente: los cambios conceptuales, y la programación de computadoras. Sin embargo, al hacerlo, tuvimos que sacrificar ciertos aspectos del estudio de la enseñanza entre pares y su interacción con variables demográficas y cognitivas, como se desarrolló en detalle en la sección “4.1 Introducción”. En este capítulo, entonces, seleccionamos un dominio no tanto por su relevancia sino por su flexibilidad para abordar el estudio de la enseñanza entre pares con mayor resolución. Así, caracterizamos un popular paradigma, identificando factores que determinan el desempeño en la tarea, y lo adaptamos, optimizando su simplicidad de administración y análisis, permitiéndonos aumentar el número de participantes, evaluaciones y variables consideradas.

En este capítulo nuevamente observamos diferencias en el desempeño según les participantes fueran enseñades por une adulte o une par, y una compensación entre precisión y tiempo de respuesta. Además, observamos que estas diferencias dependen de la edad, el género y la tendencia a pedir explicaciones aclaratorias, y la dificultad de la tarea. Nuevamente, no pudimos encontrar efectos de enseñar sobre el propio desempeño, aún después de incluir instancias adicionales de revisión, pero sí identificamos una posible modulación de la habilidad metacognitiva sobre el aprovechamiento de esta instancia de enseñanza.

Los resultados de los dos estudios presentados por un lado representan un aporte metodológico para estudiar la enseñanza entre pares con un paradigma con ventajas para su administración y análisis. Además, la consideración de múltiples variables (edad, género, dificultad de la tarea), posibilitada justamente por este paradigma, refleja en parte la complejidad de los estudios en el área de enseñanza entre pares. En suma, los resultados de este capítulo ofrecen metodologías y evidencia empírica que podrían apoyar y guiar prácticas de enseñanza entre pares y estudios adicionales en el campo.

Sin embargo, no obstante las ventajas y avances con respecto a los estudios presentados en los capítulos precedentes, quedan cuestiones pendientes, como la no verificación del efecto tutor, una mayor profundización sobre la relación con la metacognición, y la consideración de las propias instancias de enseñanza entre pares. En el último capítulo a continuación repasaremos estas cuestiones y propondremos posibles aproximaciones futuras.

Capítulo 5: Discusión final

En esta tesis nos propusimos explorar el potencial de la tutoría entre pares para el aprendizaje en dominios del conocimiento de relevancia para el mundo actual pero que permanecían sin explorar en el contexto de esta estrategia pedagógica, describiendo posibles efectos sobre otras variables cognitivas, y evaluando el impacto de distintos moduladores.

Con este fin llevamos a cabo tres estudios principales, desarrollados en detalle en cada uno de los capítulos precedentes. En estos estudios básicamente formamos parejas de estudiantes con distintos niveles de conocimiento acerca de un tema y evaluamos el impacto de la interacción sobre ambes participantes, como así también de individues control que no participaron de la tarea de interacción.

A continuación, recapitulamos los resultados de estos estudios, resumiéndolos y comparándolos a la luz de los objetivos originales del presente trabajo, discutimos el impacto que estos resultados podrían tener en el área, y comentamos limitaciones y direcciones futuras.

5.1 Aproximación metodológica

El objetivo de este trabajo fue ofrecer evidencia empírica sobre el potencial de la tutoría entre pares como estrategia pedagógica complementaria, en dominios de relevancia sociocultural que presentan complejidades para el aprendizaje. Asimismo, nos propusimos describir diferencias con otras formas de enseñanza, beneficios para todes les participantes a corto y largo plazo, y relaciones con otras habilidades cognitivas; siempre considerando variables de contexto y de organización relevantes.

Con este fin realizamos tres estudios experimentales que básicamente consistieron de una etapa de evaluación inicial, seguida de una etapa de interacción entre participantes con distintos niveles de conocimiento (une tutore y une aprendiz) o de una etapa control de revisión a solas, y una etapa de evaluación final. Además, en algunos casos se incluyó una etapa de instrucción previa para una parte de les participantes con el objeto de generar una diferencia de conocimiento (*brecha*

artificial) entre tutores y aprendices, mientras que en otros casos se aprovecharon las diferencias preexistentes (*brecha natural*) en las aulas.

De este modo pudimos comparar: (1) aprender de une par *vs.* aprender de une adulte, y en algunos casos también *vs.* aprender a solas, sin instrucción; y (2) revisar el material enseñando a une par *vs.* hacerlo a solas. Concretamente evaluamos desempeño y en algunos casos metacognición también, ambos en el contexto del dominio de conocimiento tratado. El desempeño fue evaluado como precisión y tiempo de respuesta por separado, y en algunos casos con una variable que combina a ambos. La metacognición fue evaluada de forma directa, o indirectamente como la confianza en las respuestas dadas.

Los estudios de esta tesis fueron desarrollados consecutivamente como se describió en detalle en los capítulos precedentes, permitiéndonos aprovechar los aprendizajes de cada uno para mejorar el diseño de los siguientes.

5.2 Resumen de los estudios

En esta sección resumimos las características y resultados de los estudios principales 1 a 3, correspondientes a los Capítulos 2 a 4, respectivamente. A su vez, estos son resumidos y comparados en las tablas 5.1, 5.2 y 5.3.

	Estudio 1	Estudio 2	Estudio 3
Dominio	Cambios conceptuales	Lenguajes de programación	Reglas jerárquicas
Interacción	Tutoría entre pares (implícita)	Tutoría entre pares	Tutoría entre pares
Brecha de conocimiento	Natural	Artificial	Artificial
Tamaño muestral	42	42	124
Rango etario (años)	7-8	7-8	7-12
Reevaluación aprendiz	No	No	Sí
Evaluación diferida	No	Sí	No
Aprender de par vs de adulto	No	Sí	Sí
Aprender de par vs sin instrucción	Sí	Sí	No
Enseñar a par vs revisar solo	Sí	Sí	Sí

Tabla 5.1: *Diferencias entre estudios principales*

Estudio	**Grupo etario**	**n**		
		Control	**Tutor**	**Aprendiz**
1	7-8	12	15	15
2*	7-8	12	15	15
3	7-8	10	13	13
	9-10	21	18	18
	11-12	9	11	11
	Total	**64**	**72**	**72**

Tabla 5.2: *Distribución de casos entre estudios y grupos experimentales*

* A fin de incluir también un control sin instrucción, en el estudio 2 se realizó un estudio complementario del que participaron además 11 niñes enseñades por une adulte y 9 niñes sin instrucción de entre 7 y 10 años. Los detalles en el Capítulo 3.

5.2.1 Cambios conceptuales

En el *Estudio principal 1* del Capítulo 2 explotamos la diversidad de modelos mentales del planeta Tierra que esperábamos encontrar en les participantes y evaluamos el efecto de interacción entre niñes con distintos niveles de conocimiento.

Utilizando una tarea de aprendizaje asistido por pares con asimetría implícita de roles y conocimiento, semejante a una tutoría entre pares, comparamos (1) aprender de une par *vs.* aprender a solas, y (2) enseñar a une par *vs.* revisar a solas.

En primer lugar, confirmamos la diversidad de modelos que esperábamos encontrar según trabajos previos. Luego, observamos que la interacción entre pares redujo la brecha natural de nivel de conocimiento entre elles. Concretamente, identificamos que esto fue producto del avance de les niñes menos avanzades, a través de la incorporación de respuestas de sus pares más avanzades, resultando en procesos de cambio conceptual sobre aspectos relativamente simples. Por otra parte, no observamos mejoras en el nivel de conocimiento de les estudiantes más avanzades, aunque no observamos deterioros tampoco.

Cabe destacar que en este capítulo también presentamos un posible método alternativo para simplificar la evaluación de modelos mentales usando mapas conceptuales.

5.2.2 Lenguajes de programación

En el *Estudio principal 2* del Capítulo 3, a diferencia del estudio anterior, introdujimos una brecha de conocimiento artificial, enseñando aspectos básicos de programación a una parte de les participantes. A continuación, algunes de elles debían enseñar a compañeres que no habían participado de la clase inicial. Asimismo, incluimos una etapa de evaluación diferida a dos años, y un control adicional sin instrucción. De este modo pudimos comparar (1) aprender de une par *vs.* aprender de une adulte *vs.* aprender a solas (sin instrucción), y (2) enseñar a une par *vs.* revisar a solas.

En principio, aprender de une par resultó indistinguible de aprender de une adulte, considerando al desempeño como una combinación de precisión y velocidad. Sin embargo, al considerarlas de forma separada encontramos diferencias. Concretamente, observamos que aprender de une par resultó en respuestas más rápidas pero menos precisas que aprender de une adulte (aunque más precisas que aprender a solas). Estas diferencias desaparecieron en la evaluación diferida dos años después. De forma análoga, lo que parecían diferencias de desempeño entre participantes de distinto género, resultaron ser diferencias de velocidad de resolución, cuando consideramos precisión y velocidad por separado.

Por otra parte, no observamos mayor desempeño en quienes enseñaron a une par en comparación con quienes revisaron a solas, aunque nuevamente no observamos un perjuicio tampoco, ni en la evaluación inmediata ni en la diferida.

Por último, a pesar de un mayor desempeño en la evaluación final, quienes enseñaron a une par no se mostraron más confiades de sus respuestas, a diferencia de quienes no enseñaron. Además, acompañando la caída de desempeño en la evaluación diferida, tutores y aprendices se mostraron menos confiades, mientras que quienes no participaron de la tutoría entre pares mantuvieron niveles de confianza altos. Esto fue interpretado como una posible modulación de la metacognición mediada por tutoría entre pares.

5.2.3 Reglas jerárquicas

Finalmente, en el *Estudio principal 3* del Capítulo 4 utilizamos nuevamente un protocolo que incluía una brecha artificial, esta vez enseñando a una parte de les participantes las reglas de un juego de percepción visual. En este caso, a diferencia de los estudios anteriores, no elegimos el dominio por su relevancia sociocultural, sino por su facilidad de evaluación y análisis, lo que permitió ampliar la muestra, el número de evaluaciones y las variables consideradas. Comparamos aprender de une par *vs.* aprender de une adulte, y enseñar a une par *vs.* revisar a solas (ampliando el número de participantes en este último grupo al sumar una etapa de reevaluación para aprendices también).

En primer lugar, observamos diferencias en el equilibrio precisión-velocidad. Similar al resultado obtenido en el estudio anterior, quienes aprendieron de une par mostraron mayor velocidad a costas de menor precisión, aunque en este caso esto dependió de la edad, el género, la confianza y la dificultad de la tarea.

Por otra parte, en líneas generales en este estudio tampoco encontramos beneficios de enseñar sobre el propio desempeño. No obstante, identificamos una posible modulación por habilidad metacognitiva del aprovechamiento que les tutores hacen de la etapa de enseñanza, aunque este resultado resta ser confirmado.

Finalmente, en este estudio no encontramos diferencias en la confianza sobre el propio desempeño entre ninguna de las condiciones.

Cabe destacar que para llevar a cabo este estudio caracterizamos un popular paradigma y lo adaptamos para el estudio experimental masivo de la enseñanza entre pares, ofreciendo de este modo un aporte metodológico para futuras investigaciones.

	Estudio	1 (Cap. 2)	2 (Cap. 3)		3 (Cap. 4)
	Evaluación	Inmediata	Inmediata	Diferida	Inmediata
Basal	Desempeño Agregado	n.a.	*Gén:* Mayor en niños		n.d.(2)
	Desempeño Precisión	Reprodujimos diversidad descripta previamente	*Gén:* Sin diferencias entre géneros		*Edd:* Crece con la edad *Gén:* Mayor en niños
	Desempeño Velocidad	n.a.	*Gén:* Mayor en niños (proporción de ejercicios finalizados)		Sin diferencias por edad o género
	Metacognición	n.a.	n.d.		Media similar a trabajos previos; *Edd:* Crece con edad; *Gén:* Mayor en niños
Aprender de une par (Aprendiz) vs aprender de un adulto	Desempeño Agregado		Sin diferencias	Sin diferencias	n.d.(2)
	Desempeño Precisión	n.a.	Menor precisión	Sin diferencias	Menor; *Edd:* Sólo 9-12 años; *Dif:* Sólo dificultad intermedia; *Gén:* Sólo niños; *Met:* Sólo alta confianza
	Desempeño Velocidad		Mayor velocidad	n.d.	Mayor velocidad
	Metacognición		Sin diferencias de confianza	Sin diferencias (vs Tutor); Menor confianza (vs Ctrl.)	Sin diferencias de confianza
Aprender de une par (Aprendiz) vs sin instrucción	Desempeño Agregado	n.a.	n.a.(1)		n.a.
	Desempeño Precisión	Mayor; *Dif:* Concentrado en aspectos de dificultad baja	Mayor precisión	Sin diferencias	
	Desempeño Velocidad	n.a.	n.d.	n.d.	
	Metacognición	n.a.	n.d.	n.d.	
Enseñar a une par (Tutor) vs revisar a solas (Control)	Desempeño Agregado	n.a.	Sin diferencias	Sin diferencias	n.d.(2)
	Desempeño Precisión	Sin diferencias; *Brc:* Posible modulación por brecha con aprendiz	Sin diferencias	Sin diferencias	Sin diferencias; *Met:* Posible modulación por habilidad metacognitiva (3)
	Desempeño Velocidad	n.a.	Sin diferencias	n.d.	n.d.(4)
	Metacognición	n.a.	Menor confianza	Menor confianza	Sin diferencias de confianza

Tabla 5.3: *Resumen de resultados*

En algunos casos se incluye modulación por: metacognición (*Met)*, dificultad de la tarea (*Dif*), brecha tutor-aprendiz (*Brc*), género (*Gén*), o edad (*Edd*).

n.a. (no aplica): No es posible evaluar este aspecto con el protocolo empleado; n.d. (no disponible): Este aspecto no fue analizado.

(1) En el *Estudio principal 2* (Cap. 3), el desempeño agregado en la condición *sin-instructor* no es comparable con el de la condición *par-instructor* debido a diferencias entre las sesiones (ver apartado *Condición sin-instructor* de la sección "3.2.2.1 Efecto del tipo de instrucción").

(2) Ver comentario sobre *eficiencia inversa* en la sección "4.3.3 Discusión".

(3) Resultados descriptivos, no significativos.

(4) En el *Estudio principal 3* (Cap. 4), no se cumplen los supuestos de las pruebas paramétricas para comparar tiempos de respuesta entre *controles* y *tutores*; preferimos enfocarnos en la precisión como medida del desempeño.

5.3 Avances reportados

A continuación consideramos los resultados resumidos arriba, esta vez en contexto específico de los objetivos e hipótesis presentados en el capítulo 1, comentando los avances que estos resultados reportan para el estudio de la tutoría entre pares.

5.3.1 Dominios de relevancia

Por un lado nuestros estudios aportan evidencia empírica a favor del uso de la tutoría entre pares en dominios específicos del conocimiento de relevancia para el contexto sociocultural actual, en particular el de los cambios conceptuales y el de la programación de computadoras. Específicamente, demuestran que la interacción entre pares con niveles de conocimiento diferentes en estos dominios reduce la brecha de conocimiento entre ambes, ya sea que esta brecha sea el resultado de diferencias preexistentes o generadas artificialmente.

De esta manera nuestro trabajo amplía el conjunto de dominios del conocimiento en los que se ha explorado el potencial de la tutoría entre pares, siendo en particular nuestro estudio sobre programación de computadoras el primero en explorar la cuestión de la tutoría entre pares en niñes para este dominio.

5.3.2 Rol pedagógico complementario

En todos los estudios mostramos que les estudiantes pueden aprender de sus pares, pudiendo incluso sugerir en algunos casos que en primera instancia parecen aprender tanto como de une adulte. Sin embargo, análisis más detallados nos permitieron identificar diferencias entre aprender de une par y aprender de une adulte, subrayando la importancia de considerar múltiples ángulos de una misma variable. En particular identificamos una compensación entre precisión y tiempo de respuesta, observando en general que les que aprenden de une par son más rápidos, a costas de una precisión un poco menor.

Es posible que la presencia de une par instructore, a diferencia de la figura de autoridad que significa une instructore adulte, haya sesgado el desempeño de les niñes, de una forma cuidadosa que evita errores hacia una forma más relajada, quizá más lúdica, que tolera más riesgos por prueba y error (Higgings, 1997), promoviendo un comportamiento exploratorio. En este sentido, la tutoría entre pares podría haber afectado las nociones que les niñes tienen acerca de cómo utilizar los nuevos conocimientos, sugiriendo una suerte de interacción entre esta estrategia pedagógica y sus creencias epistémicas (Oh et al., 2013).

Estos resultados resaltan el rol complementario de la tutoría entre pares, en tanto complemento y no sustituto de otras formas de enseñanza. En este sentido, las diferencias descriptas podrían guiar decisiones sobre cuándo usar esta forma de enseñanza en lugar de otras.

5.3.3 Efecto tutor

A pesar de los beneficios que esperábamos encontrar sobre el desempeño de les que enseñaron a une par como ha sido descripto en trabajos previos, no encontramos efecto de la tutoría entre pares sobre el aprendizaje de les tutores en ninguno de los estudios realizados.

Algunas posibles razones detrás de estos resultados fueron descriptas en los capítulos correspondientes a cada estudio. Por ejemplo, en el *Estudio principal 1* del Capítulo 2 identificamos un efecto techo en la medida de desempeño que podía estar enmascarando aprendizajes de les tutores, y en el *Estudio principal 2* del Capítulo 3 postulamos que podían ser necesarias instancias de revisión después de

enseñar para capitalizar sus beneficios. Ambas cuestiones fueron atendidas en los estudios subsiguientes, pero en ningún caso se observaron beneficios para les tutores.

No obstante, a pesar de no haber observado que les tutores aprendieran por enseñar como esperábamos, cabe destacar que no observamos que éstes se vieran perjudicades tampoco, un aspecto que suele despertar preocupaciones (Kalkowski, 1995), como comentamos en el Capítulo 1. Además, sí observamos posibles efectos sobre la confianza que éstes tienen acerca de su desempeño. Éstos podrían sugerir beneficios sobre la habilidad metacognitiva, como discutiremos en la sección "5.3.5.1 Efectos sobre metacognición", a continuación.

5.3.4 Efectos a largo plazo

Otra cuestión que queríamos abordar en nuestros estudios es el impacto a largo plazo de la tutoría entre pares. Este aspecto pudo ser evaluado en sólo uno de los estudios que realizamos, dos años después de la intervención. Sin embargo, les participantes de todas las condiciones mostraron el mismo desempeño, indistinguible del manifestado por participantes que no habían recibido instrucción alguna.

No obstante, sí observamos diferencias a largo plazo en la confianza que les participantes mostraron acerca de su propio desempeño, como veremos a continuación.

5.3.5 Metacognición

En este trabajo nos propusimos profundizar el estudio de la relación entre la tutoría entre pares y la metacognición, comprendida como la habilidad de monitorear y controlar el propio conocimiento (Kornell et al., 2007; Nelson & Narens, 1990). Analizamos tanto el impacto de la tutoría entre pares sobre aspectos relacionados con esta capacidad, como la modulación que la metacognición podría ejercer sobre los alcances de esta estrategia pedagógica.

5.3.5.1 Efectos sobre metacognición

Es esperable que una habilidad general como la metacognición no sería afectada inmediatamente por las intervenciones puntuales y cortas de nuestros estudios. Por

tanto, en su lugar evaluamos una variable relacionada, la confianza acerca del propio desempeño, en el contexto específico de las tareas presentadas.

Esta variable fue evaluada en dos de nuestros estudios. Como resumimos anteriormente, en uno de ellos, les niñes que participaron de la tutoría entre pares, tanto aprendices como tutores, mostraron menor confianza a corto y largo plazo que les que aprendieron de une adulte y revisaron a solas. Como desarrollamos en la sección “3.2.3.3 Enseñanza entre pares y metacognición”, teniendo en cuenta que les niñes suelen mostrarse exageradamente confiades (Finn & Metcalfe, 2014), este resultado podría sugerir que la tutoría entre pares habría modulado las habilidades metacognitivas de les participantes, atenuando así sus elevados niveles de confianza. No obstante, esta cuestión debería ser considerada con mayor detenimiento (ver sección “5.4.3 Metacognición”, más adelante).

Además, este resultado no pudo ser reproducido en el otro estudio en que consideramos la variable confianza, en el que no observamos diferencias entre condiciones.

5.3.5.2 Modulación por metacognición

Por otra parte, nos interesaba también considerar la relación opuesta. Es decir, cómo la habilidad metacognitiva afecta el desarrollo y los alcances de la tutoría entre pares.

Por un lado, en el *Estudio principal 3* vimos que participantes que confiaban menos en sus respuestas (quizá indicativo de una mejor metacognición) y por tanto pedían explicaciones más frecuentemente no mostraban la característica menor precisión cuando eran enseñades por une par.

Por otro lado, en ese mismo estudio, parecía ser que en algunos casos una mayor habilidad metacognitiva iba a acompañada por mejor aprendizaje por parte de les tutores. Sin embargo, este resultado no es estadísticamente significativo y sólo puede guiarnos en el diseño de futuros estudios que busquen confirmar esta posible relación.

5.3.6 Variables de contexto y organización

Como desarrollamos en el Capítulo 1, la tutoría entre pares es un conjunto algo heterogéneo de prácticas caracterizadas por una larga serie de variables organizacionales, y que también dependen de una serie de variables de contexto. Por esta razón, en nuestros estudios tuvimos en cuenta algunas de estas variables en el diseño de las intervenciones y en el análisis de los resultados:

- **Edad:** Si bien en dos estudios tuvimos en cuenta la edad como covariable o factor de los análisis realizados, sólo en el *Estudio principal 3* las diferencias de edad fueron lo suficientemente amplias como para observar su influencia. Concretamente, en este estudio observamos que el desempeño de les participantes más jóvenes no era tan bueno como para notar diferencias entre aprender de une par y aprender de une adulte. No obstante, estas diferencias sí fueron observables en el *Estudio principal 2* para el mismo grupo etario. Esto resalta la importancia de la edad (y la dificultad de la tarea, como veremos a continuación) a la hora de considerar el uso de esta estrategia pedagógica.
- **Género:** En uno de los estudios (*Estudio principal 3*) observamos que las diferencias entre aprender de une par y aprender de une adulte se manifestaron solamente en participantes de género masculino. Este resultado no fue reproducido en el otro estudio (*Estudio principal 2*) en que comparamos estos dos tipos de instrucción, aunque en ese caso trabajamos con edades diferentes. Por otra parte, gracias a nuestra aproximación metodológica, pudimos indagar en diferencias de desempeño entre géneros descriptas en trabajos previos y mostrar que al menos en nuestro caso se deberían a diferencias de velocidad y no de precisión.
- **Amplitud de la brecha:** En el *Estudio principal 1*, en que aprovechamos diferencias de conocimiento preexistentes en vez de generarlas artificialmente, evaluamos la amplitud de la brecha de conocimiento entre les miembros de las parejas como posible modulador de los resultados obtenidos. Así observamos que distancias menores con le participante menos avanzade parecen haber resultado en mejor aprendizaje en le participante

más avanzade. Sin embargo, como discutimos en el capítulo correspondiente, este resultado no es robusto y futuros estudios deberían confirmarlo.

- **Dificultad de la tarea:** Observamos también que algunos efectos dependieron de la dificultad de la tarea evaluada. Por ejemplo, en el *Estudio principal 1* el aprendizaje de les estudiantes menos avanzades habría girado en torno a la revisión de los modelos mentales alternativos más simples, y en el *Estudio principal 3* las diferencias entre aprender de une par y aprender de une adulte se habrían manifestado en ejercicios de dificultad intermedia solamente. Estos resultados subrayan la importancia de considerar tareas de dificultad diversa para estimar el impacto de la tutoría entre pares, ya que ciertos efectos podrían pasar desapercibidos en tareas demasiado sencillas o demasiado complejas.

5.3.7 Aportes metodológicos

Por último, si bien realizar aportes metodológicos al campo de estudio de la tutoría entre pares no era uno de los objetivos originales de este trabajo, nos gustaría destacar algunos de estos aportes que, creemos, surgen de nuestros estudios.

En primer lugar, en el Capítulo 2 proponemos una forma alternativa, más rápida y sencilla, de evaluar modelos mentales haciendo uso de mapas conceptuales construidos automáticamente a partir del discurso oral o escrito de les participantes. Por otra parte, en el estudio principal del Capítulo 3 presentamos una formalización del lenguaje de programación utilizado, que podría servir para otros estudios que hagan uso de éste. Finalmente, en el Capítulo 4 ofrecemos una caracterización de un paradigma interesante para el estudio de la tutoría entre pares, por sus reglas complejas y simplicidad de evaluación y análisis, y una adaptación para su distribución masiva y controlada.

5.4 Direcciones futuras

Nuestros resultados reportan avances en el área de estudio de las tutorías entre pares. Sin embargo, al mismo tiempo, abren nuevas preguntas. Además, no todas las cuestiones han podido ser exploradas en suficiente profundidad y otras han quedado sin tratar. En esta sección comentamos sobre algunos de estos aspectos,

sugiriendo algunos caminos de estudio futuros en relación a la tutoría entre pares en niñes.

5.4.1 Equilibrio precisión-velocidad

Algunos de nuestros resultados sugieren que la tutoría entre pares podría diferenciarse de otras formas de enseñanza en la relación que les aprendices tienen con el conocimiento que obtienen de ésta, favoreciendo un uso más exploratorio que prioriza los tiempos de respuesta por sobre la precisión.

¿A qué se debe esta diferencia? ¿Acaso quienes aprenden de une par creen estar jugando y les que aprenden de une adulte estar siendo evaluades y esto cambia la forma en la que encaran la tarea? De ser así, ¿podríamos afectar los resultados explicitando la naturaleza lúdica (o evaluativa) de la tarea? ¿O acaso la mayor velocidad es simple consecuencia de una menor habilidad?

Estas cuestiones, de relevancia a la hora de elegir usar la tutoría entre pares entre otras estrategias de enseñanza en función de los resultados deseados, deberán ser abordadas en estudios futuros en el área.

5.4.2 Efecto tutor

Como desarrollamos en el Capítulo 1, la tutoría entre pares ha demostrado beneficiar no sólo a les aprendices, sino a les tutores también. Sin embargo, no observamos este *efecto tutor* en ninguno de nuestros estudios.

Entendemos que algunas decisiones de diseño de nuestras investigaciones habrían sido las adecuadas para favorecer este efecto. Por ejemplo, la naturaleza estructurada provista por los materiales de instrucción, la pequeña diferencia de edad entre les participantes, y la conformación de parejas del mismo género (Leung, 2019). Sin embargo, algunas decisiones podrían haber sido subóptimas, como veremos a continuación.

En primer lugar, es posible que la falta de entrenamiento para les tutores les haya sesgado a formas poco productivas de recitar conocimiento o *knowledge-telling* (Roscoe & Chi, 2007). Por otra parte, si bien la estructura aportada por los materiales de instrucción podría reportar ventajas generales (K. Topping, 2017), y específicas para les tutores (Leung, 2019), es posible también que éstos hayan

sesgado a les tutores a estas formas improductivas de recitar conocimiento, impidiéndoles además aprender mediante un esfuerzo de recuperación o *learning by retrieval* (Koh et al., 2018).

En segundo lugar, por razones logísticas nuestros estudios involucraban una única sesión de tutoría entre pares por participante. Sin embargo, habría evidencia de que sesiones más numerosas y prolongadas favorecerían el efecto tutor (Leung, 2019). Además, futuros estudios que incluyan una etapa de instrucción previa (como en nuestros estudios principales 2 y 3) podrían aprovechar también el efecto que la expectativa de enseñar ejercería sobre el aprendizaje (Bargh & Schul, 1980), como fue desarrollado en el Capítulo 1, avisando a les participantes, antes de esta etapa de instrucción, que deberán enseñar.

Por otra parte, también es posible que los efectos de enseñar vs revisar a solas se manifestaran a largo plazo (Hermida et al., 2021). Si bien consideramos esta posibilidad, sólo pudimos evaluarla en uno de los estudios (ver sección “3.2.2.3 Efectos a largo plazo” del *Estudio principal 2*), tras un plazo quizá demasiado largo. Futuros estudios podrían abordar esta cuestión en mayor detalle.

Por último, otras variables a tener en cuenta para promover el *efecto tutor* podrían incluir la brecha de conocimiento entre aprendices y tutores, y los niveles de metacognición. Algunos resultados de nuestros estudios sugerirían que brechas menores y mayor capacidad metacognitiva podrían estar asociados con mayores aprendizajes por parte de les tutores. Sin embargo, estos resultados son preliminares y carecen de solidez estadística, por lo que sólo pueden inspirar estudios futuros que busquen confirmarlos.

5.4.3 Metacognición

En el último estudio evaluamos metacognición de forma directa, empleando una tarea especializada. Esta medida resultó útil para considerar el rol modulador de esta habilidad sobre el impacto de la tutoría entre pares. Sin embargo, obtenerla no fue sencillo, pues implicaba administrar una evaluación adicional a les participantes. Además, su utilidad para estudiar la relación inversa, es decir el efecto que nuestras intervenciones podrían tener sobre la metacognición, es limitada, pues era de

esperar que nuestras intervenciones puntuales y cortas no afectaran la maduración de esta habilidad cognitiva general.

Por estos motivos evaluamos también el uso de estas habilidades en el contexto específico de las tareas administradas, de forma indirecta a través de la confianza que les participantes reportaban respecto de su desempeño en dichas tareas. Esto nos permitió identificar posibles efectos de la tutoría entre pares sobre estos reportes de confianza, y postular interpretaciones en términos de habilidades metacognitivas.

Sin embargo, usar estos reportes de confianza como medida de metacognición tiene sus limitaciones. En particular, es necesario tener suficientes datos que permitan observar reportes de confianza en ejercicios respondidos correcta e incorrectamente. Sin embargo, esto no estaba disponible en nuestros estudios, que no fueron diseñados con este objetivo en mente. Por lo tanto, nuestras interpretaciones de los resultados de confianza en términos de metacognición son especulativas y dependen de suposiciones, como la discutida en el Capítulo 3 acerca de la excesiva confianza que acostumbrarían mostrar les niñes (Finn & Metcalfe, 2014).

Por lo tanto, para entender mejor cómo la tutoría entre pares podría afectar el uso de habilidades metacognitivas en contexto del tema enseñado y contrastar nuestras interpretaciones, será necesario realizar estudios que incluyan tareas diseñadas específicamente con este objetivo, que controlen el desempeño para garantizar suficientes ejercicios respondidos correcta e incorrectamente, con alta y baja confianza. Un objetivo diferente es indagar cómo la tutoría entre pares afecta la metacognición de forma general y transferible a otras tareas. Es posible que en este caso sean necesarias intervenciones más prolongadas que las sesiones puntuales y cortas empleadas en nuestros estudios.

Por último, retomando la cuestión de la evaluación directa de la metacognición, mencionada al comienzo de esta sección, cabe mencionar algunas limitaciones también. Puntualmente, como comentamos en el capítulo correspondiente, a diferencia de estudios anteriores esta evaluación se hizo de forma autoadministrada. Si bien la metacognición media obtenida fue similar a la reportada en estudios previos, encontramos diferencias por edad y género que no esperábamos. Es posible que estas diferencias sean consecuencia de la forma de administración y

que estén reflejando aspectos no relacionados con la metacognición. Por lo tanto, estudios futuros que busquen seguir indagando acerca del rol modulador de la metacognición sobre los efectos de la tutoría entre pares, deberán mejorar la administración de la tarea correspondiente para evitar posibles artefactos.

5.4.4 Interacciones

Por otra parte, estudiar las interacciones entre participantes es clave para identificar los rasgos comportamentales que caracterizan a la tutoría entre pares y determinan su efectividad. Si bien en algunos de nuestros estudios pudimos registrarlas en detalle, decidimos priorizar otros análisis en el tiempo disponible, por lo que acabamos considerándolas como una suerte de caja negra.

Futuros estudios que se enfoquen específicamente en estas interacciones podrán ampliar nuestra comprensión acerca de cómo y por qué la tutoría entre pares funciona, para mejorar su efectividad.

5.4.5 Validez ecológica y generalización

Entendemos por validez ecológica al grado en que los resultados experimentales son extrapolables a situaciones análogas del mundo “real”. Los estudios presentados en esta tesis fueron experimentales, permitiendo tener mayor control sobre las variables y resultados más robustos. Sin embargo, esta aproximación limita el número de participantes y la validez ecológica, como veremos a continuación.

Por un lado, el estudio 1 aprovechó diferencias de conocimiento preexistentes en el aula para formar parejas de participantes con distintos niveles. Esta estrategia es más ecológica que la empleada en los estudios 2 y 3, en los que se incluyó una etapa de instrucción previa para una parte de les participantes en cada clase; algo que en un escenario real podría ser más complejo de implementar.

No obstante, el estudio 1 también se caracterizó por entrevistas une a une para evaluar los niveles de conocimiento. Esta estrategia es probablemente poco aplicable al aula. Por tal motivo, al final del capítulo correspondiente proponemos una posible alternativa para simplificar y quizá automatizar la evaluación de estos modelos mentales. Por el contrario, los estudios 2 y 3 emplearon formas

autoadministradas de evaluación, mucho más trasladables a contextos reales. En este sentido, el estudio 3 reúne condiciones similares a las que podríamos encontrar en un aula real, con muches participantes interactuando o siendo evaluados al mismo tiempo en un mismo espacio.

Por otra parte, cabe mencionar que todos los estudios incluyen en la misma clase una condición experimental con interacción y otra control sin interacción. Esta concurrencia de condiciones en una misma clase podría ser problemática de extrapolar a situaciones reales.

Estudios futuros de la tutoría entre pares en los dominios tratados en esta tesis podrían maximizar la validez ecológica combinando estrategias de nuestros estudios. Por ejemplo, podrían incluir una etapa de instrucción previa generalizada, que evite tener que dividir la clase. Luego una etapa de evaluación inicial autoadminstrada para estimar el nivel de conocimiento de cada participante y formar parejas, como en el estudio 1. A continuación una etapa de interacción simultánea, como en el estudio 3, y finalmente una evaluación final autoadministrada. Para evitar el problema de condiciones con y sin interacción en una misma clase, el grupo control sin interacción podría sustituirse por uno con interacción acerca de un tema no relacionado. Alternativamente, las condiciones experimental y control podrían asignarse a clases diferentes, aunque esta aproximación presenta otras limitaciones.

En cualquier caso, es importante que tales estudios tengan en cuenta la forma de asignar condiciones y parejas a les participantes. En nuestros estudios esto se hizo de forma pseudoaleatoria, asegurando que variables como género y nivel de conocimiento estuvieran balanceadas entre condiciones. Sin embargo, es posible que en escenarios menos controlados se pongan en juego otros factores que podrían sesgar estas decisiones.

Por último, en cuanto a la generalización de nuestros resultados, cabe recordar que la población de nuestros tres estudios principales es relativamente homogénea, limitándose a estudiantes de escuela primaria de nivel socioeconómico medio-alto y alto. Por su parte, otros grupos de trabajo en nuestro laboratorio, trabajando con poblaciones más heterogéneas, han encontrado diferencias en el impacto de sus intervenciones en función de la tasa de asistencia a clase por razones familiares de les participantes (Goldin et al., 2014). Además, sabemos que el impacto de la tutoría

entre pares sería mayor en poblaciones con bajo desempeño académico (Bowman-Perrott et al., 2013). Por lo tanto, repetir nuestros estudios en otras poblaciones podría ser importante a fin de generalizar los resultados obtenidos.

5.5 Conclusión final

En esta tesis exploramos la tutoría entre pares en niñes como estrategia pedagógica para apoyar el aprendizaje en dominios de relevancia sociocultural, incluyendo su interacción con variables potencialmente relacionadas. En esta sección final comentaremos los posibles alcances educativos y socioculturales de nuestro trabajo.

5.5.1 Educación

Nuestros resultados aportan evidencia empírica a favor del uso de esta estrategia para la enseñanza en dominios de relevancia que presentan dificultades para el aprendizaje. Asimismo, subrayan su rol complementario, identificando diferencias entre aprender de une par y aprender de une adulte. Estas diferencias podrán servir de guía para decidir en qué circunstancias emplear esta práctica frente a otras estrategias de enseñanza, ya sea en el aula o en el diseño de políticas educativas.

Por otra parte, la evidencia, si bien parcial, de la posible relación entre la tutoría entre pares y la metacognición podría alentar la exploración de esta práctica no sólo como herramienta para promover el aprendizaje, sino también para la consolidación de esta importante habilidad cognitiva.

Por último, nuestros resultados sugieren que los temores acerca del perjuicio que la tutoría entre pares podría tener sobre les estudiantes más avanzades estarían infundados. No obstante, nos recuerdan que los beneficios para éstes no son automáticos, sino que dependen del ajuste adecuado de múltiples variables.

5.5.2 Cultura y sociedad

Nuestro trabajo ofrece evidencia a favor de una forma de apoyar el aprendizaje en aspectos relevantes para el contexto sociocultural actual. Por un lado, en el actual contexto de crisis de posverdad, desinformación y negacionismo científico, mostramos que niñes con distintos conceptos del planeta Tierra son capaces de llegar a un consenso en dirección hacia el conocimiento científicamente aceptado,

promoviendo procesos de cambio conceptual en les participantes menos avanzades. Por otra parte, en un contexto crecientemente dominado por tecnologías digitales, mostramos que la tutoría entre pares puede apoyar el aprendizaje de lenguajes de programación en niñes, ofreciendo una herramienta para el empoderamiento tecnológico de generaciones futuras.

Por otra parte, más allá del rol como facilitadora del aprendizaje, la tutoría entre pares podría tener un impacto sociocultural independiente del dominio específico abordado. Las propuestas curriculares que reconocen el conocimiento de les estudiantes y que crean oportunidades para que éstes construyan con otres influirían en sus creencias epistémicas (Hofer, 2001); es decir, sus nociones e intuiciones acerca de qué es el conocimiento, cómo surge, cómo se transmite, a quién pertenece, etc. La tutoría entre pares en particular, como un entorno de aprendizaje constructivista, tiene el potencial de promover la sofisticación de estas creencias (Oh et al., 2013), como también sugieren algunos de nuestros resultados.

Orientar futuras investigaciones en esta dirección podría subrayar el rol de la tutoría entre pares como una herramienta no sólo para reforzar el aprendizaje, sino también para fomentar prácticas productivas de construcción colaborativa y promover otras formas de relación con el conocimiento. En el mundo actual, en que cada vez más las fuentes de autoridad se desdibujan y las personas podemos alzar nuestras voces y ejercer nuestra influencia de forma inmediata alrededor del planeta, quizá estas formas nos ayuden a guiar estos procesos de cambio en pos del fortalecimiento de nuestros tejidos sociales para el beneficio de la humanidad.

Referencias

Allan, N. P., Lonigan, C. J., & Wilson, S. B. (2013). Psychometric evaluation of the Children's Behavior Questionnaire-Very Short Form in preschool children using parent and teacher report. *Early Childhood Research Quarterly*, *28*(2), 302–313. https://doi.org/10.1016/j.ecresq.2012.07.009

Altintas, T., Gunes, A., & Sayan, H. (2016). A peer-assisted learning experience in computer programming language learning and developing computer programming skills. *Innovations in Education and Teaching International*, *53*(3), 329–337. https://doi.org/10.1080/14703297.2014.993418

Altszyler, E., Ribeiro, S., Sigman, M., & Fernández Slezak, D. (2017). The interpretation of dream meaning: Resolving ambiguity using Latent Semantic Analysis in a small corpus of text. *Consciousness and Cognition*, *56*, 178–187. https://doi.org/10.1016/j.concog.2017.09.004

Asterhan, C. S. C., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, *99*(3), 626–639. https://doi.org/10.1037/0022-0663.99.3.626

Asterhan, C. S. C., & Schwarz, B. B. (2009). Argumentation and Explanation in Conceptual Change: Indications From Protocol Analyses of Peer-to-Peer Dialog. *Cognitive Science*, *33*(3), 374–400. https://doi.org/10.1111/j.1551-6709.2009.01017.x

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific reasoning. *Social Development*, *2*(3), 202–221. https://doi.org/10.1111/j.1467-9507.1993.tb00014.x

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science (New York, N.Y.)*, *329*(5995), 1081–1085. https://doi.org/10.1126/science.1185718

Balanskat, A., & Engelhardt, K. (2015). *Computing our future. Computer programming and coding: Priorities, school curricula and initiatives across Europe [Technical report]*. European Schoolnet. http://www.eun.org/resources/detail?publicationID=661

Bär, N. (2018, febrero 2). *Charlatanes y excéntricos*. La Nación. https://www.lanacion.com.ar/opinion/charlatanes-y-excentricos-nid2105928

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology*, *72*(5), 593.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860–1863. https://doi.org/10.1073/pnas.0910967107

Bell, A. (1807). *An analysis of the experiment in education, made at Egmore, near Madras ...* Printed by T. Bensley ... for Cadell and Davies.

Benotti, L., Martínez, M. C., & Schapachnik, F. (2014). Engaging high school students using chatbots. *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education - ITiCSE '14*, 63–68. https://doi.org/10.1145/2591708.2591728

Bian, L., Leslie, S.-J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389–391. https://doi.org/10.1126/science.aah6524

Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, *13*(6), 4–16.

Blown, E. J. (2013). Children's Concepts of the Shape and Size of the Earth, Sun and Moon. *International Journal of Science Education*, *35*(3), 388–446. https://doi.org/10.1080/09500693.2012.750432

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous

exploration and discovery. *Cognition*, *120*(3), 322–330. https://doi.org/10.1016/j.cognition.2010.10.001

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer Tutoring: A meta-analytic review of single-case research. *School Psychology Review*, *42*(1), 39–55. https://doi.org/10.1016/j.jcomdis.2015.06.009

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *108 Suppl*, 10918–10925. https://doi.org/10.1073/pnas.1100290108

Braught, G., Wahls, T., & Eby, L. M. (2011). The Case for Pair Programming in the Computer Science Classroom. *ACM Transactions on Computing Education*, *11*(1), 1–21. https://doi.org/10.1145/1921607.1921609

Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2019). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-019-0748-6

Brinda, T., Puhlmann, H., & Schulte, C. (2009). Bridging ICT and CS: Educational Standards for Computer Science in Lower Secondary Education. *Proceedings of the 14th Annual ACM SIGCSE Conference on Innovation and Technology in Computer Science Education*, 288–292. https://doi.org/10.1145/1562877.1562965

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press.

Brown, N. C. C., Kölling, M., Crick, T., Peyton Jones, S., Humphreys, S., & Sentance, S. (2013). Bringing Computer Science Back into Schools: Lessons from the UK. *Proceeding of the 44th ACM Technical Symposium on Computer Science Education*, 269–274. https://doi.org/10.1145/2445196.2445277

Bruyer, R., & Brysbaert, M. (2011). Combining Speed and Accuracy in Cognitive Psychology: Is the Inverse Efficiency Score (IES) a Better Dependent Variable

than the Mean Reaction Time (RT) and the Percentage Of Errors (PE)? *Psychologica Belgica*, *51*(1), 5–13. https://doi.org/10.5334/pb-51-1-5

Buchs, C., Butera, F., Mugny, G., & Darnon, C. (2004). Conflict Elaboration and Cognitive Outcomes. *Theory Into Practice*, *43*(1), 23–30. https://doi.org/10.1207/s15430421tip4301_4

Butera, F., Darnon, C., & Mugny, G. (2011). Learning from Conflict. En *Rebels in Groups* (pp. 36–53). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444390841.ch3

Caldwell, C. A., & Millen, A. E. (2009). Social Learning Mechanisms and Cumulative Cultural Evolution: Is Imitation Necessary? *Psychological Science*, *20*(12), 1478–1483.

Calero, C. I. (2020). ¿Qué saben, ¿qué dicen que saben y qué enseñan los niños? *Revista Argentina de Ciencias del Comportamiento (RACC)*, *12*(Extra 1 (Suplemento I (Mayo)), 38–39.

Calero, C. I., Goldin, A. P., & Sigman, M. (2018). The Teaching Instinct. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-018-0383-6

Calero, C. I., Shalom, D. E., Spelke, E. S., & Sigman, M. (2019). Language, gesture, and judgment: Children's paths to abstract geometry. *Journal of Experimental Child Psychology*, *177*, 70–85. https://doi.org/10.1016/j.jecp.2018.07.015

Calero, C. I., Zylberberg, A., Ais, J., Semelman, M., & Sigman, M. (2015). Young children are natural pedagogues. *Cognitive Development*, *35*, 65–78. https://doi.org/10.1016/j.cogdev.2015.03.001

Carey, S. (1991). Knowledge acquisition: Enrichment or conceptual change? En S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 257–291). Lawrence Erlbaum Associates.

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press.

Carey, S., & Spelke, E. S. (1994). Domain-specific knowledge and conceptual change. En L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169–200). Cambridge University Press.

Carlson, S. M., Koenig, M. A., & Harms, M. B. (2013). Theory of mind. *WIREs Cognitive Science*, *4*(4), 391–402. https://doi.org/10.1002/wcs.1232

Caro, T. M., & Hauser, M. D. (1992). Is there teaching in nonhuman animals? *The Quarterly review of biology*, *67*(2), 151–174.

Cassel, W. S., Roebers, C. E. M., & Bjorklund, D. F. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggestive questions. *Journal of Experimental Child Psychology*, *61*, 116–133. http://dx.doi.org/10.1006/jecp.1996.0008

Castellaro, M., & Roselli, N. (2018). Interacción sociocognitiva entre pares en situaciones simétricas y asimétricas de competencia epistémica. *Revista de Psicología*, *36*(1), 333–365. https://doi.org/10.18800/psico.201801.011

Chi, M. T. H., De Leeuw, N., Chiu, M. H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439–477. https://doi.org/10.1016/0364-0213(94)90016-7

Cialdini, R. B. (2006). *Influence: The Psychology of Persuasion, Revised Edition* (Revised edition). Harper Business.

Cohen, J. (1986). Theoretical considerations of peer tutoring. *Psychology in the Schools*, *23*(2), 175–186. https://doi.org/10.1002/1520-6807(198604)23:2<175::AID-PITS2310230211>3.0.CO;2-H

Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational Outcomes of Tutoring: A Meta-Analysis of Findings. *American Educational Research Journal*, *19*(2), 237–248.

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The Effect of Instructional Explanations on Learning from Scientific Texts. *The Journal of the Learning Sciences*, *6*(4), 347–365.

Consejo Federal de Educación de Argentina. (2015). *Resolución 263/15*.

Consejo Federal de Educación de Argentina. (2018). *Resolución 343/18*. http://www.bnm.me.gov.ar/giga1/normas/RCFE_343-18.pdf

Corriveau, K. H., & Winters, M. A. (2019). Trusting Your Teacher: Implications for Policy. *Policy Insights from the Behavioral and Brain Sciences*, *6*(2), 123–129. https://doi.org/10.1177/2372732219862574

Crawford, T. (2005). What counts as knowing: Constructing a communicative repertoire for student demonstration of knowledge in science. *Journal of Research in Science Teaching*, *42*(2), 139–165. https://doi.org/10.1002/tea.20047

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in cognitive sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Curtis, D. D., & Lawson, M. J. (2019). Exploring collaborative online learning. *Online Learning*, *5*(1), Article 1. https://doi.org/10.24059/olj.v5i1.1885

Dalbey, J., & Linn, M. C. (1985). The Demands and Requirements of Computer Programming: A Literature Review. *Journal of Educational Computing Research*, *1*(3), 253–274. https://doi.org/10.2190/BC76-8479-YM0X-7FUA

Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research*, *13*(1), 9–19. https://doi.org/10.1016/0883-0355(89)90013-X

Dangwal, R., & Kapur, P. (2009). Learning through teaching: Peer-mediated instruction in minimally invasive education. *British Journal of Educational Technology*, *40*(1), 5–22. https://doi.org/10.1111/j.1467-8535.2008.00863.x

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to Scientific Theories Affects Women's Math Performance. *Science*, *314*(5798), 435–435. https://doi.org/10.1126/science.1131100

Davis, B. L., & Maclagan, D. (2003). The card game SET. *The Mathematical Intelligencer*, *25*(3), 33–40.

De Backer, L., Van Keer, H., & Valcke, M. (2012). Exploring the potential impact of reciprocal peer tutoring on higher education students' metacognitive knowledge and regulation. *Instructional Science*, *40*(3), 559–588. https://doi.org/10.1007/s11251-011-9190-5

de la Hera, D. P., Sigman, M., & Calero, C. I. (2019). Social interaction and conceptual change pave the way away from children's misconceptions about the Earth. *npj Science of Learning*, *4*(1). https://doi.org/10.1038/s41539-019-0051-3

Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the Social and Cognitive Processes Underlying Human Cumulative Culture. *Science*, *335*(6072), 1114–1118. https://doi.org/10.1126/science.1213969

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417–440. https://doi.org/10/dwxmnj

Dijk, J. A. van. (2010). *The network society: Social aspects of new media* (2. ed., reprinted). SAGE.

diSessa, A. A. (2009). A Bird's-Eye View of the "Pieces" vs. "Coherence" Controversy (from the "Pieces" side of the Fence). En S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 35–60). Routledge.

Duran, D. (2016). Learning-by-teaching. Evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International*, *3297*(March), 1–9. https://doi.org/10.1080/14703297.2016.1156011

Eckstein, M. K., Starr, A., & Bunge, S. A. (2019). How the inference of hierarchical rules unfolds over time. *Cognition*, *185*, 151–162. https://doi.org/10.1016/j.cognition.2019.01.009

Edwards, L. D., Coddington, A., & Caterina, D. (1997). Girls teach themselves, and boys too: Peer learning in a computer-based design and construction activity. *Computers & Education*, *29*(1), 33–48. https://doi.org/10.1016/S0360-1315(97)00018-3

Eikospentaki, K., Tsonos, D., Kouroupetroglou, G., & Vosniadou, S. (2015). The Development of Knowledge about the Earth and the Day/Night Cycle in Blind and Sighted Children Using Acoustical Rendition of Documents' Visual

Elements. *Procedia Computer Science*, *65*(October), 484–491. https://doi.org/10.1016/j.procs.2015.09.119

Ellis, J. A., Semb, G. B., & Cole, B. (1998). Very Long-Term Memory for Information Taught in School. *Contemporary Educational Psychology*, *23*(4), 419–433. https://doi.org/10.1006/ceps.1997.0976

Fedorenko, E., Ivanova, A., Dhamala, R., & Bers, M. U. (2019). The Language of Programming: A Cognitive Perspective. *Trends in Cognitive Sciences*, *23*(7), 525–528. https://doi.org/10.1016/j.tics.2019.04.010

Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, *407*(6804), 630–633. https://doi.org/10/cwndrg

Fessakis, G., Gouli, E., & Mavroudi, E. (2013). Problem solving by 5–6 years old kindergarten children in a computer programming environment: A case study. *Computers & Education*, *63*, 87–97. https://doi.org/10.1016/j.compedu.2012.11.016

Finn, B., & Metcalfe, J. (2014). Overconfidence in children's multi-trial judgments of learning. *Learning and Instruction*, *32*, 1–9. https://doi.org/10.1016/j.learninstruc.2014.01.001

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Flannery, L. P., Kazakoff, E. R., Bontá, P., Silverman, B., Bers, M. U., & Resnick, M. (2013). Designing ScratchJr: Support for Early Childhood Learning Through Computer Programming. *Proceedings of the 12th International Conference on Interaction Design and Children (IDC '13)*, 1–10. https://doi.org/10.1145/2485760.2485785

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *0*. https://doi.org/10.3389/fnhum.2014.00443

Frède, V., Nobes, G., Frappart, S., Panagiotaki, G., Troadec, B., & Martin, A. (2011). The acquisition of scientific knowledge: The influence of methods of

questioning and analysis on the interpretation of children's conceptions of the earth. *Infant and Child Development*, *20*(6), 432–448. https://doi.org/10.1002/icd.730

Fundación Sadosky. (2016). *Una propuesta para refundar la enseñanza de la computación en las escuelas Argentinas* (CC–2016). Fundación Sadosky.

Galton, F. (1907). Vox Populi. *Nature*, *75*, 450–451.

Ghetti, S., Hembacher, E., & Coughlin, C. A. (2013). Feeling uncertain and acting on it during the preschool years: A metacognitive approach. *Child Development Perspectives*, *7*(3), 160–165. https://doi.org/10.1111/cdep.12035

Goldin, A. P., Hermida, M. J., Shalom, D. E., Elias Costa, M., Lopez-Rosenfeld, M., Segretin, M. S., Fernandez-Slezak, D., Lipina, S. J., & Sigman, M. (2014). Far transfer to language and math of a short software-based gaming intervention. *Proceedings of the National Academy of Sciences*, *111*(17), 6443–6448. https://doi.org/10.1073/pnas.1320217111

Golding, P., Facey-Shaw, L., & Tennant, V. (2006). Effects of Peer Tutoring, Attitude and Personality on Academic Performance of First Year Introductory Programming Students. *Proceedings. Frontiers in Education. 36th Annual Conference*, 7–12. https://doi.org/10.1109/FIE.2006.322662

Goodlad, S., & Hirst, B. (1989). *Peer Tutoring. A Guide to Learning by Teaching.* ERIC.

Gopnik, A., Meltzoff, A. N., & Bryant, P. (1997). *Words, thoughts, and theories* (Vol. 1). Mit Press Cambridge, MA.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, *9*(6), 495–522. https://doi.org/10.1002/acp.2350090604

Gregory, A., Walker, I., Mclaughlin, K., & Peets, A. D. (2011). Both preparing to teach and teaching positively impact learning outcomes for peer teachers. *Medical Teacher*, *33*(8), e417–e422.

Grover, S., & Pea, R. (2013). Computational Thinking in K–12: A Review of the State of the Field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Gülbahar, Y., & Kalelioğlu, F. (2014). The Effects of Teaching Programming via Scratch on Problem Solving Skills: A Discussion from Learners' Perspective. *Informatics in Education - An International Journal*, *13*(1), 33–50.

Hadzigeorgiou, Y. (2015). Young Children's Ideas About Physical Science Concepts. En K. Cabe Trundle & M. Saçkes (Eds.), *Research in Early Childhood Science Education* (pp. 67–97). Springer Netherlands. http://link.springer.com/10.1007/978-94-017-9505-0_4

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed). Psychology Press.

Harvey, J., & Hays, D. G. (1972). Effect of Dogmatism and Authority of the Source of Communication upon Persuasion. *Psychological Reports*, *30*(1), 119–122. https://doi.org/10.2466/pr0.1972.30.1.119

Hermida, M. J., Santangelo, A. P., Calero, C. I., Goizueta, C., Espinosa, M., & Sigman, M. (2021). Learning-by-Teaching Approach Improves Dengue Knowledge in Children and Parents. *The American Journal of Tropical Medicine and Hygiene*, *1*(aop). https://doi.org/10.4269/ajtmh.21-0253

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*(3), 141–158. https://doi.org/10.1119/1.2343497

Heyd-Metzuyanim, E., & Schwarz, B. B. (2017). Conceptual change within dyadic interactions: The dance of conceptual and material agency. *Instructional Science*, *45*(5), 645–677. https://doi.org/10.1007/s11251-017-9419-z

Higgings, E. T. (1997). Beyond pleasure and pain. *American psychologist*, *52*(12), 1280.

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. AAUW.

Hobson, S. M., Trundle, K. C., & Saçkes, M. (2010). Using a Planetarium Software Program to Promote Conceptual Change with Young Children. *Journal of Science Education and Technology*, *19*(2), 165–176. https://doi.org/10.1007/s10956-009-9189-8

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational psychology review*, *13*(4), 353–383.

Hoppitt, W., Brown, G., Kendal, R., Rendell, L., Thornton, A., Webster, M., & Laland, K. (2008). Lessons from animal teaching. *Trends in Ecology & Evolution*, *23*(9), 486–493. https://doi.org/10.1016/j.tree.2008.05.008

Howe, C., Tolmie, A., & Rodgers, C. (1990). Physics in the primary school: Peer interaction and the understanding of floating and sinking. *European Journal of Psychology of Education*, *5*(4), 459–475. https://doi.org/10.1007/BF03173132

Howe, N., & Recchia, H. (2005). Playmates and teachers: Reciprocal and complementary interactions between siblings. *Journal of family psychology : JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *19*(4), 497–502. https://doi.org/10.1037/0893-3200.19.4.497

Hunt, A., & Thomas, D. (1999). *The Pragmatic Programmer: From Journeyman to Master*. Addison Wesley.

Jacob, M., & Hochstein, S. (2008). Set recognition as a window to perceptual and cognitive processes. *Perception & Psychophysics*, *70*(7), 1165–1184. https://doi.org/10.3758/PP.70.7.1165

Jenkins, J. R., & Jenkins, L. M. (1987). Making Peer Tutoring Work. *Educational Leadership*, *44*(6), 64.

Kalelioğlu, F. (2015). A new way of teaching programming skills to K-12 students: Code.org. *Computers in Human Behavior*, *52*, 200–210. https://doi.org/10.1016/j.chb.2015.05.047

Kalkowski, P. (1995). Peer and Cross-Age Tutoring. En *Series IX, 1994-95* (p. 25). School Improvement Program of the Northwest Regional Educational Laboratory.

Kata, A. (2010). A postmodern Pandora's box: Anti-vaccination misinformation on the Internet. *Vaccine*, *28*(7), 1709–1716. https://doi.org/10.1016/j.vaccine.2009.12.022

Kazakoff, E. R., Sullivan, A., & Bers, M. U. (2013). The Effect of a Classroom-Based Intensive Robotics and Programming Workshop on Sequencing Ability in Early Childhood. *Early Childhood Education Journal*, *41*(4), 245–255. https://doi.org/10.1007/s10643-012-0554-5

Keil, F. C. (1994). *Concepts, kinds, and cognitive development*. MIT Press.

Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. *International Journal of Science Education*, *20*(7), 849–871. https://doi.org/10.1080/0950069980200707

King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Review*, *10*(1), 57–74. https://doi.org/10.1023/A:1022858115001

Kline, M. (2014). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences*, 1–71. https://doi.org/10.1017/S0140525X14000090

Koh, A. W. L., Lee, S. C., & Lim, S. W. H. (2018). The learning benefits of teaching: A retrieval practice hypothesis. *Applied Cognitive Psychology*, *32*(3), 401–410. https://doi.org/10.1002/acp.3410

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological science*, *18*(1), 64–71. https://doi.org/10.1111/j.1467-9280.2007.01850.x

Kuhn, D. (2015). Thinking Together and Alone. *Educational Researcher*, *44*(1), 46–53. https://doi.org/10.3102/0013189X15569530

Kuhn, T. S. (1957). *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Harvard University Press.

Kuhn, T. S. (1971). *The structure of scientific revolutions*. University Press.

Kurzban, R., & Barrett, H. C. (2012). Origins of Cumulative Culture. *Science*, *335*(6072), 1056–1057. https://doi.org/10.1126/science.1219232

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review*, *10*, 133–149. https://doi.org/10.1016/j.edurev.2013.02.002

Laal, M., & Laal, M. (2012). Collaborative learning: What is it? *Procedia - Social and Behavioral Sciences*, *31*, 491–495. https://doi.org/10.1016/j.sbspro.2011.12.092

Landrum, A. (2019, febrero 17). *YouTube as a Primary Propagator of Flat Earth Philosophy*. American Association for the Advancement of Science (AAAS). http://www.asheleylandrum.com/alternative-beliefs-project.html

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

LeBlanc, G., & Bearison, D. J. (2004). Teaching and learning as a bi-directional activity: Investigating dyadic interactions between child teachers and child learners. *Cognitive Development*, *19*(4), 499–515. https://doi.org/10.1016/j.cogdev.2004.09.004

Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology*, *107*(2), 558–579. https://doi.org/10.1037/a0037698

Leung, K. C. (2019). An updated meta-analysis on the effect of peer tutoring on tutors' achievement. *School Psychology International*, *40*(2), 200–214. https://doi.org/10.1177/0143034318808832

Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady*, *10*, 707.

Levin, H. M., Glass, G. V., & Meister, G. R. (1987). Cost-Effectiveness of Computer-Assisted Instruction. *Evaluation Review*, *11*(1), 50–72. https://doi.org/10.1177/0193841X8701100103

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction*, *11*(4–5), 357–380. https://doi.org/10.1016/S0959-4752(00)00037-2

Littlejohn, S., & Foss, K. (2009). Social Construction of Reality. En *Encyclopedia of Communication Theory* (Vol. 1–2, pp. 892–894). SAGE Publications, Inc. https://doi.org/10.4135/9781412959384

Llaurado, A., Martí, M. A., & Tolchinsky, L. (2012). Corpus CesCa: Compiling a corpus of written Catalan produced by school children. *International Journal of Corpus Linguistics*, *17*(3), 428–441. https://doi.org/10.1075/ijcl.17.3.06lla

Lochhead, J., & Whimbey, A. (1987). Teaching analytical reasoning through thinking aloud pair problem solving. *New Directions for Teaching and Learning*, *1987*(30), 73–92. https://doi.org/10.1002/tl.37219873007

Lombrozo, T. (2006). The structure and function of explanations. *Trends in cognitive sciences*, *10*(10), 464–470. https://doi.org/10.1016/j.tics.2006.08.004

López, C., & Narodowski, M. (1999). El mejor de los métodos posibles: La introducción del método lancasteriano en Iberoamérica en el temprano siglo XIX. En M. H. C. Bastos, L. M. de F. Filho, & P. Lesage (Eds.), *A escola elementar no século XIX: o método monitorial/mútuo*. Universidade de Passo Fundo, EDIUPF.

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, *108*(22), 9020–9025. https://doi.org/10.1073/pnas.1008636108

Lui, K. M., & Chan, K. C. C. (2006). Pair programming productivity: Novice-novice vs. Expert-expert. *International Journal of Human Computer Studies*, *64*(9), 915–925. https://doi.org/10.1016/j.ijhcs.2006.04.010

Mackay, C. (1841). *Extraordinary Popular Delusions the Madness of Crowds*. Wordsworth Editions Limited.

Mali, G. B., & Howe, A. (1979). Development of earth and gravity concepts among Nepali children. *Science Education*, *63*(5), 685–691. https://doi.org/10.1002/sce.3730630514

Maloney, J., Resnick, M., & Rusk, N. (2010). The Scratch programming language and environment. *ACM Transactions on Computing Education*, *10*(4), 1–15. https://doi.org/10.1145/1868358.1868363.http

Marno, H., Danyi, R., Vékony, T., Janacsek, K., & Németh, D. (2021). Prompting teaching modulates children's encoding of novel information by facilitating higher-level structure learning and hindering lower-level statistical learning. *Cognition*, *213*, 104784. https://doi.org/10.1016/j.cognition.2021.104784

Martinez, C., Gomez, M. J., & Benotti, L. (2015). A Comparison of Preschool and Elementary School Children Learning Computer Science Concepts Through a Multilanguage Robot Programming Platform. *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, 159–164. https://doi.org/10.1145/2729094.2742599

Mayerová, K., & Veselovská, M. (2014). The Programming Environment for the Lego WeDo Robotic Construction Set. En J. Kapounová & K. Kostolányová (Eds.), *Information and Communication Technology in Education* (pp. 149–157). University of Ostrava. https://core.ac.uk/download/pdf/33693352.pdf#page=149

Mayerové, K., & Veselovská, M. (2017). How to Teach with LEGO WeDo at Primary School. En M. Merdan, W. Lepuschitz, G. Koppensteiner, & R. Balogh (Eds.), *Robotics in Education* (Vol. 457, pp. 55–62). Springer International Publishing. https://doi.org/10.1007/978-3-319-42975-5_5

McClelland, M. M., Geldhof, G. J., Cameron, C. E., & Wanless, S. B. (2015). Development and Self-Regulation. En *Handbook of Child Psychology and Developmental Science* (pp. 1–43). American Cancer Society. https://doi.org/10.1002/9781118963418.childpsy114

Mead, G. H., Morris, C. W., Huebner, D. R., & Joas, H. (2015). *Mind, self, and society* (The definitive edition). University of Chicago Press.

Mercier, H. (2011). Reasoning serves argumentation in children. *Cognitive Development*, *26*(3), 177–191. https://doi.org/10.1016/j.cogdev.2010.12.001

Mercier, H., & Sperber, D. (2012). "Two Heads Are Better" Stands to Reason. *Science*, *336*(6084), 979–979. https://doi.org/10.1126/science.336.6084.979-a

Miller, D., Topping, K., & Thurston, A. (2010). Peer tutoring in reading: The effects of role and organization on two dimensions of self-esteem. *The British journal of educational psychology*, *80*(Pt 3), 417–433. https://doi.org/10.1348/000709909X481652

Miller, S. A., & Brownell, C. A. (1975). Peers, Persuasion, and Piaget: Dyadic Interaction between Conservers and Nonconservers. *Child Development*, *46*(4), 992–997. https://doi.org/10.2307/1128411

Mitra, S., Dangwal, R., Chatterjee, S., Jha, S., Bisht, R. S., & Kapur, P. (2005). Acquisition of computing literacy on shared public computers: Children and the "hole in the wall". *Australasian Journal of Educational Technology*, *21*(3). https://doi.org/10.14742/ajet.1328

Mohan, M. (1972). *Peer tutoring as a technique for teaching the unmotivated.* State University of New York, Teacher Education Research Center. http://eric.ed.gov/?id=ED061154

Moshakis, A. (2018, mayo 27). *Is the Earth flat? Meet the people questioning science*. The Guardian. https://www.theguardian.com/global/2018/may/27/is-the-earth-pancake-flat-among-the-flat-earthers-conspiracy-theories-fake-news

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology*, *8*(2), 181–192. https://doi.org/10.1002/ejsp.2420080204

Navajas, J., Niella, T., Garbulsky, G., Bahrami, B., & Sigman, M. (2018). Aggregated knowledge from a small number of debates outperforms the wisdom of large crowds. *Nature Human Behaviour*, *2*(2), 126–132. https://doi.org/10.1038/s41562-017-0273-4

Nelson, T. O., & Narens, L. (1990). Metamemory: A Theoretical Framework and New Findings. En G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 26, pp. 125–173). Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5

Nestojko, J. F., Bui, D. C., Kornell, N., & Bjork, E. L. (2014). Expecting to teach enhances learning and organization of knowledge in free recall of text passages. *Memory & Cognition*, *42*(7), 1038–1048. https://doi.org/10.3758/s13421-014-0416-z

Nosek, J. T. (1998). The case for collaborative programming. *Communications of the ACM*, *41*(3), 105–108. https://doi.org/10.1145/272287.272333

Nourbakhsh, I. R., Hamner, E., Crowley, K., & Wilkinson, K. (2004). Formal measures of learning in a secondary school mobile robotics course. *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, 1831-1836 Vol.2. https://doi.org/10.1109/ROBOT.2004.1308090

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology*, *33*(3), 345–359. https://doi.org/10.1016/j.cedpsych.2008.06.001

Nussbaum, J. (1979). Children's conceptions of the earth as a cosmic body: A cross age study. *Science Education*, *63*(1), 83–93. https://doi.org/10.1002/sce.3730630113

Oh, S. A., Kim, Y. S., Lee, S.-D., & Kim, N.-R. (2013). *The Effect of Peer Tutoring in Terms of Epistemological Beliefs and Self-Regulated Learning*. 48. http://icome2013.iwd.jp/program/pdf/1p_PDF/A38.pdf

OPSSI. (2019). *Reporte anual sobre el Sector de Software y Servicios Informáticos de la República Argentina* (Núm. 2018). Cámara de Empresas de Software y Servicios Informáticos de la República Argentina. https://www.cessi.org.ar/descarga-institucionales-2330/documento2-290a5d80ccf2da8c5d3935209aa10568

O′Reilly, D. (1998). School programming as literacy: The case for BOXER. *Journal of Computer Assisted Learning*, *14*(1), 51–58. https://doi.org/10.1046/j.1365-2729.1998.1410051.x

Osborne, J. (2010). Arguing to Learn in Science: The Role of Collaborative, Critical Discourse. *Science*, *328*(5977), 463–466. https://doi.org/10.1126/science.1183944

Padró, L., & Stanilovsky, E. (2012, mayo). FreeLing 3.0: Towards Wider Multilinguality. *Proceedings of the Language Resources and Evaluation Conference (LREC 2012)*.

Panagiotaki, G., Nobes, G., & Potton, A. (2009). Mental models and other misconceptions in children's understanding of the earth. *Journal of Experimental Child Psychology*, *104*(1), 52–67. https://doi.org/10.1016/j.jecp.2008.10.003

Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, Inc.

Parera, L. (2019, febrero 6). *Insólito. Dicen que la Tierra es plana y se reúnen este fin de semana en Buenos Aires*. La Nación.

https://www.lanacion.com.ar/sociedad/municipio-bonaerense-ayuda-organizar-encuentro-creen-nid2223626

Perry, M., Breckinridge Church, R., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*(4), 359–400. https://doi.org/10.1016/0885-2014(88)90021-4

Phillips, R., Lockton, D., Baurley, S., & Silve, S. (2013). Making instructions for others: Exploring mental models through a simple exercise. *Interactions*, *20*(5), 74–79. https://doi.org/10.1145/2505290

Piaget, J. (1929). *The child's conception of the world*. Routledge & Kegan Paul.

Piaget, J. (1932). *The moral judgment of the child*. Routledge.

Piaget, J. (1950). *The Psychology of Intelligence*. Routledge & Kegan Paul.

Piaget, J. (1985). *The equilibration of cognitive structures*. University of Chicago Press.

Pinto-Llorente, A. M., Casillas-Martín, S., Cabezas-González, M., & García-Peñalvo, F. J. (2018). Building, coding and programming 3D models via a visual programming environment. *Quality & Quantity*, *52*(6), 2455–2468. https://doi.org/10.1007/s11135-017-0509-4

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The behavioral and brain sciences*, *4*, 515–526. https://doi.org/10.1016/j.celrep.2011.1011.1001.7.

President's Council of Advisors on Science and Technology. (2010). *Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology. Report to the President and Congress*. Executive Office of the President. https://eric.ed.gov/?id=ED527261

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of personality assessment*, *87*(1), 102–112. https://doi.org/10.1207/s15327752jpa8701_09

Raafat, R. M., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in Cognitive Sciences*, *13*(10), 420–428. https://doi.org/10.1016/j.tics.2009.08.002

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Resnick, M., Silverman, B., Kafai, Y., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., & Silver, J. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60. https://doi.org/10.1145/1592761.1592779

Rideout, V. (2013). Zero to eight: Children's media use in America 2013. *Pridobljeno*, *11*(1).

Rippon, G. (2019). Chapter 12: Good Girls Don't. En *Gender and Our Brains: How New Neuroscience Explodes the Myths of the Male and Female Minds*.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, *95*(2), 240–257. https://doi.org/10.1037/0022-0663.95.2.240

Roscoe, R. D. (2014). Self-monitoring and knowledge-building in learning by teaching. *Instructional Science*, *42*(3), 327–351. https://doi.org/10.1007/s11251-013-9283-4

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding Tutor Learning: Knowledge-Building and Knowledge-Telling in Peer Tutors' Explanations and Questions. *Review of Educational Research*, *77*(X), 534–574. https://doi.org/10.3102/0034654307309920

Roscoe, R. D., & Chi, M. T. H. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science*, *36*, 321–350. https://doi.org/10.1007/s11251-007-9034-5

Rothbart, M. K., Ahadi, S. a, Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child development*, *72*(5), 1394–1408. https://doi.org/10.1111/1467-8624.00355

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning Words from Knowledgeable versus Ignorant Speakers: Links Between Preschoolers' Theory of Mind and Semantic Development. *Child Development*, *72*(4), 1054–1070. https://doi.org/10.1111/1467-8624.00334

Sabol, M. A., & Wisher, R. A. (2001). Retention and Reacquisition of Military Skills. *Military Operations Research*, *6*(1), 59–80. JSTOR.

Saçkes, M. (2015). Young Children's Ideas About Earth and Space Science Concepts. En K. Cabe Trundle & M. Saçkes (Eds.), *Research in Early Childhood Science Education* (pp. 35–65). Springer Netherlands. http://link.springer.com/10.1007/978-94-017-9505-0_3

Salinas, C. G., García, M. A., Carnicero, J. A. C., & Montesinos, M. D. H. (2000). Elaboración de una adaptación a población española del cuestionario infant behavior questionnaire para la medida del temperamento en la infancia. *Psicothema*, *12*(4), 513–519.

Salles, A., Ais, J., Semelman, M., Sigman, M., & Calero, C. I. (2016). The metacognitive abilities of children and adults. *Cognitive Development*, *40*, 101–110. https://doi.org/10.1016/j.cogdev.2016.08.009

Scherer, R. (2016). Learning from the Past–The Need for Empirical Evidence on the Transfer Effects of Computer Programming Skills. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01390

Scherer, R., Siddiq, F., & Sánchez Viveros, B. (2018). The cognitive benefits of learning computer programming: A meta-analysis of transfer effects. *Journal of Educational Psychology*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/edu0000314

Semb, G. B., Ellis, J. A., & Araujo, J. (1993). Long-term memory for knowledge learned in school. *Journal of Educational Psychology*, *85(2)*, 305–316.

Shamir, A., & Tzuriel, D. (2004). Children‘s Mediational Teaching Style as a Function of Intervention for Cross-Age Peer-Mediation. *School Psychology International*, *25*(1), 59–78. https://doi.org/10.1177/0143034304024782

Shamir, A., Zion, M., & Spector_Levi, O. (2008). Peer Tutoring, Metacognitive Processes and Multimedia Problem-based Learning: The Effect of Mediation Training on Critical Thinking. *Journal of Science Education and Technology*, *17*(4), 384–398. https://doi.org/10.1007/s10956-008-9108-4

Shi, F., Teplitskiy, M., Duede, E., & Evans, J. A. (2019). The wisdom of polarized crowds. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-019-0541-6

Siegal, M. (1991). *Knowing children: Experiments in conversation and cognition*. Erlbaum.

Slezak, D. F., & Sigman, M. (2012). Do not fear your opponent: Suboptimal changes of a prevention strategy when facing stronger opponents. *Journal of Experimental Psychology: General*, *141*(3), 527–538. https://doi.org/10.1037/a0025761

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science*, *323*(5910), 122–124. https://doi.org/10.1126/science.1165919

Sneider, G., & Pulos, S. (1983). Children's cosmographies: Understanding the earth's shape and gravity. *Science Education*, *67*(2), 205–221. https://doi.org/10.1002/sce.3730670209

Söderberg, J. (2012). *Hacking Capitalism: The Free and Open Source Software Movement* (Edición: 1). Routledge.

Soukoreff, R. W., & MacKenzie, I. S. (2001). *Measuring Errors in Text Entry Tasks: An Application of the Levenshtein String Distance Statistic*. 3.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origin of Knowledge. *Psychological Review*, *99*, 605–632. https://doi.org/10.1037/0033-295X.99.4.605

Sperber, D., & Wilson, D. (1996). *Relevance: Communication and cognition* (2nd ed). Wiley-Blackwell.

Standage, D., Wang, D.-H., Heitz, R. P., & Simen, P. (2015). Toward a unified view of the speed-accuracy trade-off. *Frontiers in Neuroscience*, *9*. https://doi.org/10.3389/fnins.2015.00139

Stöber, G. (2016). *Exploring Communicative Skills Through Peer Interactions: Young Children's Pointing and Gaze Following with Peer and Adult Partners*. Universität Leipzig.

Strauss, S., Calero, C. I., & Sigman, M. (2014). Teaching, naturally. *Trends in Neuroscience and Education*, *3*(2), 38–43. https://doi.org/10.1016/j.tine.2014.05.001

Strauss, S., & Ziv, M. (2012). Teaching is a natural cognitive ability for humans. *Mind, Brain, and Education*, *6*(4), 186–196. https://doi.org/10.1111/j.1751-228X.2012.01156.x

Strauss, S., Ziv, M., & Stein, A. (2002). Teaching as a natural cognition and its relations to preschoolers' developing theory of mind. *Cognitive Development*, *17*(3–4), 1473–1487. https://doi.org/10.1016/S0885-2014(02)00128-4

Sullivan, A., & Bers, M. U. (2013). Gender differences in kindergarteners' robotics and programming achievement. *International Journal of Technology and Design Education*, *23*(3), 691–702. https://doi.org/10.1007/s10798-012-9210-z

Thagard, P. (2009). Conceptual Change in the History of Science: Life, Mind, and Disease. En S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 374–387). Routledge.

The challenge of the post-truth era. (2018). *Nature Cell Biology*, *20*(11), 1231–1231. https://doi.org/10.1038/s41556-018-0231-z

Thornton, A., & Raihani, N. J. (2008). The evolution of teaching. *Animal Behaviour*, *75*(6), 1823–1836. https://doi.org/10.1016/J.ANBEHAV.2007.12.014

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A New Look at Infant Pointing. *Child Development*, *78*(3), 705–722. https://doi.org/10.1111/j.1467-8624.2007.01025.x

Topping, K. (2015). Peer tutoring: Old method, new developments / Tutoría entre iguales: Método antiguo, nuevos avances. *Infancia y Aprendizaje*, 1–29. https://doi.org/10.1080/02103702.2014.996407

Topping, K. (2017). *Effective peer learning: From principles to practical implementation*. Routledge, Taylor & Francis Group.

Topping, K., & Ehly, S. (1998). *Peer-assisted learning*. Routledge.

Topping, K. J. (1988). *The peer tutoring handbook: Promoting co-operative learning*. London : Croom Helm ; Cambridge, MA : Brookline Books. http://archive.org/details/peertutoringhand0000unse

Topping, K. J. (2005). Trends in Peer Learning. *Educational Psychology*, *25*(6), 631–645. https://doi.org/10.1080/01443410500345172

Topping, K. J. (2020). Peer Tutoring and Cooperative Learning. En K. J. Topping, *Oxford Research Encyclopedia of Education*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190264093.013.1432

Topping, K. J., & Ehly, S. W. (2001). Peer Assisted Learning: A Framework for Consultation. *Journal of Educational and Psychological Consultation*, *12*(2), 113–132. https://doi.org/10.1207/S1532768XJEPC1202_03

Townsend, J. T., & Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes*.

Turing, A. M. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *s2-42*(1), 230–265. https://doi.org/10/dcs37s

Tylor, E. B. (1974). *Primitive culture: Researches into the development of mythology, philosophy, religion, art, and custom*. Gordon Press.

Umapathy, K., & Ritzhaupt, A. D. (2017). A Meta-Analysis of Pair-Programming in Computer Programming Courses: Implications for Educational Practice. *ACM Transactions on Computing Education*, *17*(4), 1–13. https://doi.org/10.1145/2996201

Voogt, J., Fisser, P., Good, J., Mishra, P., & Yadav, A. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information Technologies*, *20*(4), 715–728. https://doi.org/10.1007/s10639-015-9412-6

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4*(1), 45–69. https://doi.org/10.1016/0959-4752(94)90018-3

Vosniadou, S. (2009). *International handbook of research on conceptual change*. Routledge. http://www.routledgehandbooks.com/doi/10.4324/9780203154472 https://books.google.com/books?hl=en&lr=&id=sdyOAgAAQBAJ&oi=fnd&pg=PP1&dq=international+collaboration+in+teaching+&ots=qzZ2Sqg_mj&sig=5jQtH4N-HsSEjEzAnSRxrWB6h0M

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, *24*(4), 535–585. https://doi.org/10.1016/0010-0285(92)90018-W

Vosniadou, S., & Skopeliti, I. (2014). Conceptual Change from the Framework Theory Side of the Fence. *Science & Education*, *23*(7), 1427–1445. https://doi.org/10.1007/s11191-013-9640-3

Vosniadou, S., Skopeliti, I., & Ikospentaki, K. (2005). Reconsidering the role of artifacts in reasoning: Children's understanding of the globe as a model of the earth. *Learning and Instruction*, *15*(4), 333–351. https://doi.org/10.1016/j.learninstruc.2005.07.004

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2009). The framework theory approach to the problem of conceptual change. En S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 3–34). Routledge. http://www.routledgehandbooks.com/doi/10.4324/9780203154472.ch1

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*.

Wagner, L. (1990). Social and Historical Perspectives on Peer Teaching and Education. En H. C. Foot, M. J. Morgan, & R. H. Shute (Eds.), *Children helping children* (pp. 21–42). J. Wiley.

Williams, L. (2001). Integrating pair programming into a software development process. *Proceedings 14th Conference on Software Engineering Education and Training. "In Search of a Software Engineering Profession" (Cat. No.PR01059)*, 27–36. https://doi.org/10.1109/CSEE.2001.913816

Williams, L., Kessler, R. R., Cunningham, W., & Jeffries, R. (2000). Strengthening the case for pair programming. *IEEE Software*, *17*(4), 19–25. https://doi.org/10.1109/52.854064

Williams, L., & Upchurch, R. L. (2001). In support of student pair-programming. *SIGCSE Bull.*, *33*(1), 327–331. https://doi.org/10.1145/366413.364614

Wing, J. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33–35.

Wing, J. (2017). Computational Thinking's Influence on Research and Education for All. *Italian Journal of Educational Technology*, *1*(1). https://doi.org/10.17471/2499-4324/922

Wyeth, P. (2008). How Young Children Learn to Program With Sensor, Action, and Logic Blocks. *Journal of the Learning Sciences*, *17*(4), 517–550. https://doi.org/10.1080/10508400802395069

Xu, Y., Gelfer, J. I., Sileo, N., Filler, J., & Perkins, P. G. (2008). Effects of peer tutoring on young children's social interactions. *Early Child Development and Care*, *178*(6), 617–635. https://doi.org/10.1080/03004430600857485

Yelland, N. (1993). Young Children Learning with Logo: An Analysis of Strategies and Interactions. *Journal of Educational Computing Research*, *9*(4), 465–486. https://doi.org/10.2190/B2TK-7XVG-9099-WHE7

Yuviler-Gavish, N., Faran, D., & Berman, M. (2020). The Effect of Complexity on Training for Exploration of Non-intuitive Rules in Theory of Mind. *Journal of Cognitive Enhancement*, *4*(3), 323–332. https://doi.org/10/gjq4q9

Glosario

análisis de la semántica latente. Técnica de análisis para caracterizar relaciones semánticas entre términos en un corpus.

aprendizaje social. Tipo de aprendizaje que se apoya en el comportamiento de otres.

argumentación colaborativa. Interacción en la que les individues trabajan colaborativamente para construir y criticar argumentos.

astronomía observacional. Rama de la astronomía abocada al universo observable. En el contexto del Capítulo 2, conjunto comprendido por la astronomía observacional, la cosmología y las miradas sobre el mundo.

bootstrap. Método de remuestreo, empleado en este trabajo para estimar intervalos de confianza.

cadena de texto. En informática, secuencia de caracteres.

condición adulto-instructor. En los estudios de este trabajo, condición experimental en la que une participante aprende de une adulte.

condición par-instructor. En los estudios de este trabajo, condición experimental en la que une participante aprende de une par.

conflicto sociocognitivo. Conflicto que resulta de la interacción entre individues con conocimientos u opiniones diferentes y contrastantes sobre un mismo asunto.

Constructivismo. Teoría de la educación que sostiene que el conocimiento es *construido* por quienes aprenden, en vez de *transmitido* desde alguien o algo que supuestamente lo *posee.*

distancia de Hamming. Dadas dos secuencias de la misma longitud, número de posiciones en que ambas secuencias difieren.

distancia de Levenshtein. Dadas dos secuencias de igual o distinta longitud, número de inserciones y remociones que deben hacerse a una de ellas para llegar a la otra.

elaboración epistémica. Forma de resolución del conflicto sociocognitivo que implica la consideración de argumentos y justificaciones.

elaboración relacional. Forma de resolución del conflicto sociocognitivo que considera la confianza y experiencia de les individues en conflicto.

estimación kernel de la densidad. Método para estimar la función de densidad de probabilidad de una variable.

grupo aprendiz. En los estudios de este trabajo, grupo experimental cuyes miembros aprenden de une par, y que coincide con la condición *par-instructor*.

grupo control. En los estudios de este trabajo, uno de los grupos experimentales en que es dividida la condición *adulto-instructor*, y cuyes miembros revisan a solas en lugar de enseñar a une par. Ver *grupo tutor*.

grupo tutor. En los estudios de este trabajo, uno de los grupos experimentales en que se divide la condición *adulto-instructor*, y cuyes miembros enseñan a une par, en lugar de revisar a solas. Ver *grupo control.*

hardware. En computación, partes físicas tangibles de un sistema informático. Ver *software.*

intérprete. En computación, programa que lee y ejecuta las instrucciones escritas con un lenguaje de programación en otro programa.

JavaScript. Lenguaje de programación orientado a la web.

lematización. Sustitución de una palabra o frase por su forma canónica o *lema*.

lienzo de dibujo. En programas de dibujo, área donde se inscriben los diseños (formas, trazos, etc).

navegador web. Programa de computadora que permite visitar e interactuar con páginas web.

pregunta fáctica. Aquella que puede responderse con hechos concretos, memorizables. Ver *pregunta generativa.*

pregunta generativa. Aquella cuya respuesta le sujeto debe elaborar basándose en hechos e información que conoce. Ver *pregunta fáctica.*

desempeño agregado. En los estudios de este trabajo, medidas del desempeño que combinan la precisión y el tiempo de respuesta.

Python. Popular lenguaje de programación multipropósito.

reglas jerárquicas. Sistema de reglas de decisión que comprende dos o más niveles, tal que la situación considerada debe evaluarse teniendo en cuenta todos los niveles para llegar a una conclusión.

sabiduría de las masas. Fenómeno según el cual el conocimiento colectivo de un grupo es superior al de cualquiera de sus miembros por separado.

servidor web. En computación, sistema (dispositivo o programa) que ofrece páginas web a otras partes en la misma red.

software. En computación, partes lógicas intangibles de un sistema informático. Ver *hardware.*

teoría específica. Teoría de menor jerarquía que la teoría marco en la que se inscribe, que reúne información acerca de un concepto en particular.

teoría marco. Cada una de las teorías generales en las que se agrupan y organizan los conceptos que pertenecen a un dominio del pensamiento.

word embedding. En el *análisis de la semántica latente*, representación vectorial de una palabra o frase, en el espacio semántico definido por un conjunto de textos.

z-score. Diferencia entre un valor dado y la media de su muestra o población, expresado en número de desvíos estándar de dicha muestra o población, y que permite comparar datos procedentes de muestras o poblaciones diferentes.

variable organizacional. En el contexto de tutoría entre pares, variable que describe aspectos acerca de la organización de la metodología (ej. disponibilidad de capacitaciones para tutores).

variable contextual. En el contexto de la tutoría entre pares, variable que describe aspectos del contexto en el que se desarrolla la metodología (ej. edad de les participantes).

Índice de figuras

Figura 2.1: Perspectiva del conocimiento como teoría.....25
Figura 2.2: Esquema de codificación 1, general.....36
Figura 2.3: Esquema de codificación 2, detallado.....39
Figura 2.4: Puntaje de dimensión y nivel de conocimiento.....43
Figura 2.5: Similitud de conocimiento.....44
Figura 2.6: Cambios en los niveles de conocimiento después de la interacción.....47
Figura 2.7: Cambios en la similitud de conocimiento.....50
Figura 2.8: Cambio de nivel de conocimiento como función de la brecha.....59
Figura 2.9: Distribución del número de definiciones (f) por edad.....65
Figura 3.1: Métodos.....77
Figura 3.2: Procedimiento.....83
Figura 3.3: Puntaje agregado.....92
Figura 3.4: Desempeño por ejercicio.....94
Figura 3.5: Efecto de enseñar.....100
Figura 3.6: Confianza.....102
Figura 4.1: Atributos y valores posibles.....115
Figura 4.2: Ejemplos de Set y NoSet.....116
Figura 4.3: Precisión por clase.....122
Figura 4.4: Tiempos de respuesta para ejercicios correctos.....123
Figura 4.5: Saliencia de atributos.....125
Figura 4.6: Plataforma, etapas y procedimiento.....134
Figura 4.7: Video explicativo.....137
Figura 4.8: Guía de instrucción.....138
Figura 4.9: Precisión.....148
Figura 4.10: Tiempo de respuesta.....152
Figura 4.11: Efecto de enseñar.....153

Índice de tablas

Tabla 2.1: Condiciones de análisis adicionales....52
Tabla 2.2: Resultados de análisis adicionales....53
Tabla 2.3: Cambio de nivel de conocimiento para cada grupo de temas o dominio. .56
Tabla 2.4: Puntaje inicial en Entrevista 1 por grupo de temas o dominio....56
Tabla 2.5: Diferencias significativas de distancia entre lemas seleccionados....68
Tabla 4.1: Clasificación de ejercicios....120
Tabla 4.2: Pruebas estadísticas para la saliencia de atributos....125
Tabla 4.3: Distribución por grado, clase, género y condición/grupo....131
Tabla 4.4: Edad de les participantes....132
Tabla 4.5: Distribución de ejercicios....141
Tabla 4.6: Diferencia de precisión entre condiciones por subclase de ejercicio....150
Tabla 4.7: Modulación metacognitiva del efecto de enseñar sobre el desempeño. .155
Tabla 5.1: Diferencias entre estudios principales....163
Tabla 5.2: Distribución de casos entre estudios y grupos experimentales....163
Tabla 5.3: Resumen de resultados....167

Anexos

Anexo 2.1: Cuestionario completo

Para cada pregunta se esperaban respuestas fundamentalmente verbales (V) o dibujadas (D). 30 de las preguntas eran obligatorias (O; no confundir con *temas de codificación obligada* del Esquema de codificación 2) y otras 20 estaban condicionadas (C) por respuestas previas.

Núm	Modo	Tipo	Condición	Pregunta
P01	V	C	Entrevista 1 solamente	¿Tenés hermanos? ¿Mayores, menores?
P02	D	C	Entrevista 1 solamente	Escribí tu nombre con esta birome acá.
P03	D	C	Entrevista 1 solamente	¡Muy bien! Ahora borralo, con la parte de atrás de la birome.
P04	V	O		¿Cómo se llama nuestro planeta? Planeta...
P05	V	O		¿A dónde tenemos que mirar para ver el planeta Tierra?
P06	V	O		¿Qué forma tiene el planeta Tierra?
P07	V	O		¿[forma]? ¿Cómo [forma]?
P08	V	O		Pero, ¿[forma] como qué? ¿Como qué cosa?
P09	D	O		A ver, dibujá el planeta Tierra acá.
P10	V	O		¿Así se vería desde una nave espacial?
P11	D	C	“no” en P10	¿Cómo se vería desde una nave espacial?
P12	D	O		Ahora, en el dibujo que hiciste, ¿la Luna dónde está? Dibujala.
P13	D	O		¿Y las estrellas?
P14	V	C	si dibuja luna o estrellas dentro del planeta	Entonces, la Luna/las estrellas… ¿dónde están?
P15	V	O		¿Hay más estrellas?
P16	D	C	“sí” en P16	¿Dónde hay más estrellas?
P17	D	O		¿Y el Sol? ¿Dónde está? A ver, dibujalo.
P18	D	O		¿Y el cielo?
P19	D	O		¿Y las nubes dónde están?
P20	V	C	si dibuja sol, cielo o nubes dentro del planeta	Entonces, el Sol/el cielo/las nubes ¿dónde están?
P21	V	O		¿Y hay más nubes?
P22	D	C	“sí” en P21	¿Dónde hay más nubes?
P23	V	O		¿Esto [el planeta Tierra] qué era?
P24	D	O		¡Muy bien! Ahora mostrame dónde estás vos, y dibujate ahí.
P25	D	C	si no dibuja en P24	¿Y en tu dibujo dónde estás?
P26	D	O		Ahora mostrame dónde está tu casa. Dibujala.
P27	D	O		¿Y un árbol?
P28	V	O		¿En qué país estamos?
P29	D	O		Mostrame dónde está el país.
P30	V	O		Y vos ¿estás en el país?
P31	D	C	“no” en P30	¿Dónde estás?

Núm	Modo	Tipo	Condición	Pregunta
P32	D	O		¿Y China? Mostrame dónde está.
P33	D	O		Ahora dibujá una persona que viva en China.
P34	V	O		Y esa persona, ¿puede ver el Sol, la Luna y las estrellas?
P35	V	C	“no” en P34	¿Por qué no puede verlas?
P36	D	O		¿Y el mar dónde está?
P37	V	C	si hay dibujos entre el mar y el margen inferior	¿Estos objetos están abajo del mar?
P38	D	O		Mostrame dónde te habías dibujado.
P39	D	O		Mostrame dónde llegás si caminás derecho todo el día. ¿Qué hay ahí?
P40	D	C	si no señaló el perímetro del planeta aún	Mostrame dónde llegás si seguís caminando derecho todos los días. ¿Qué hay ahí?
P41	D	C	si no señaló el perímetro del planeta aún	Mostrame dónde es lo más lejos que podés llegar. ¿Qué hay ahí?
P42	V	C	si no señaló el perímetro del planeta aún	¿Se puede llegar hasta acá [punto señalado por entrevistador]? ¿Qué hay ahí?
P43	V	O		¿Se puede seguir caminando derecho desde acá?
P44	V	O		¿Te podés caer desde acá?
P45	D	O		¿La Tierra se termina en algún lugar? ¿Tiene un borde?
P46	V	C	“sí” en P45	¿Se puede llegar hasta acá [borde]? ¿Qué hay ahí?
P47	V	C	“sí” en P45	¿Se puede seguir caminando derecho desde acá?
P48	V	C	“sí” en P45	¿Te podés caer desde acá?
P49	V	C	si no hay dibujos entre el planeta y el margen sup.	¿Qué hay acá abajo de la Tierra?
P50	V	C	si no hay dibujos entre el planeta y el margen inf.	¿Qué hay acá arriba de la Tierra?

Anexo 2.2: Guía de codificación

Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T01			**Hacia dónde mirar para ver el planeta Tierra.**	P05		
T01	0	00	Desde aquí hacia arriba o hacia el cielo.			"por un telescopio", "para arriba", "al cielo",
T01	1	10	Hacia o en el espacio.			
T01	2	20	Desde el espacio, desde afuera hacia aquí.			"desde el espacio", "desde la Luna", "hay que ir a un cohete",
T01	3	30	Desde aquí hacia abajo o hacia alrededor.			"cualquier parte", "al piso",
T02			**Forma abstracta del planeta.**	P06-08		
T02	0	00	No circular.			"cuadrada",
T02	1	10	Círculo.		Codificar gestos circulares como "círculo".	
T02	1	11	Redondo.			
T02	2	20	Esfera o casquete esférico.			sostiene una esfera en el aire,
T03			**Forma análoga al planeta.**	P06-08		
T03	0	00	Superficies cuadrangulares o infinitas.			
T03	1	10	Contornos circulares.			"como un anillo",
T03	1	11	Superficies circulares.			"como un plato",
T03	2	20	Superficies esféricas.			"como una pelota",
T04			**Dibujo del planeta como se vería desde el espacio.**	P09-11		
T04	0	00	No dibuja un círculo.			
T04	1	10	Dibuja un círculo o forma aproximadamente circular, sin continentes.			
T04	1	11	Dibuja un círculo o forma aproximadamente circular, con continentes.			
T05			**Contradicción entre planeta redondo y plano.**	P06-11		
T05	na	-01	Sin mención.			
T05	0	00	El planeta es redondo, y plano.			"el planeta es redondo, pero acá es plano",
T05	1	10	El planeta es redondo, pero se siente plano.			"como un anillo", "desde el espacio el planeta no se ve como desde adentro, sino redondo",
T06			**Forma de la Luna.**	P12	Si más de una opción posible, seleccionar la menor.	

Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T06	0	00	Irreal.			luna con rostro, forma alargada no circular ni medialuna,
T06	1	10	Medialuna.			
T06	2	20	Círculo.			
T06	2	21	Círculo con cráteres			
T07			**Ubicación de la Luna.**	P12,14	Ver figura T07. Si más de una opción posible, seleccionar la menor.	
T07	0	00	Dentro del planeta.			
T07	1	10	Abajo del planeta.			
T07	1	11	Arriba del planeta.			
T07	1	12	A los lados del planeta.			
T07	1	13	Girando alrededor del planeta.			
T08			**Múltiples posiciones de la Luna.**	P12,14		
T08	na	-01	Sin mención.			
T08	0	00	La Luna puede ocupar dos o más posiciones.		Si T07 = 13: girando alrededor del planeta, seleccionar esta opción.	
T09			**Ubicación de las estrellas.**	P13-16	Ver figura T09. Si más de una opción posible, seleccionar la menor.	
T09	0	00	Dentro del planeta.			
T09	1	10	A lo ancho de la pantalla, abajo del planeta.			
T09	1	11	A lo ancho de la pantalla, arriba del planeta.			
T09	2	20	Sólo abajo del planeta, no a lo ancho.			
T09	2	21	Sólo arriba del planeta, no a lo ancho.			
T09	2	22	Sólo a los lados del planeta.			
T09	3	30	Inmediatamente alrededor del planeta.		Inmediatamente a continuación del perímetro. No necesariamente rodeando todo el planeta.	
T09	3	31	En espacio alrededor del planeta o por todas partes.		Alrededor del planeta, no sólo abajo, arriba, a los lados o inmediatamente.	“por todas partes”,
T10			**Asociación entre estrellas y Luna y/o noche.**	P12-16	Ver figura T10.	
T10	na	-01	Sin asociación.			

Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T10	0	00	Asocia la ubicación de las estrellas con la posición de la Luna y/o de la noche.			"las estrellas siguen a la Luna", "las estrellas del lado en el que está la Luna", "las estrellas donde está la noche",
T11			**Ubicación del Sol.**	P17,20	Ver figura T11.	
T11	0	00	Dentro del planeta.			
T11	1	10	Cuarto de Sol en esquina superior izquierda o derecha.			
T11	2	20	Girando alrededor del planeta.			
T11	3	30	Abajo del planeta.			
T11	3	31	Arriba del planeta.			
T11	3	32	A los lados o atrás del planeta.			
T11	3	33	En el centro de la órbita terrestre.			"la Tierra gira alrededor del Sol",
T12			**Múltiples posiciones del Sol.**	P17,20		
T12	na	-01	Sin mención.			
T12	0	00	El Sol puede ocupar dos o más posiciones.		Si T11 = 23: girando alrededor del planeta, seleccionar esta opción.	
T13			**Ubicación del Sol relativa a la Luna.**	P12,14,17,20	Ver figura T13. Las líneas medias vertical y horizontal dividen al planeta en partes iguales.	
T13	0	00	Sin cruzar líneas medias.			
T13	0	01	Cruzando sólo línea media horizontal.			
T13	0	02	Cruzando sólo línea media vertical.			"el Sol está del otro lado de la Luna";
T13	0	03	Cruzando líneas medias vertical y horizontal.			
T14			**Representación del cielo.**	P18,20		
T14	0	00	Representa el cielo con una línea.		Gestos referidos a la pantalla se codifican en "dibujos": si traza una línea con el dedo, seleccionar esta opción.	
T14	0	01	Dibuja nubes espontáneamente.			
T14	0	02	No dibuja cielo.		Diferenciar de "no disponible" (pregunta omitida, no respondida, o evidentemente mal interpretada).	"se dibuja como está", "ya está dibujado", "no se puede dibujar",
T14	0	03	Representa el cielo con un área			
T15			**Ubicación del cielo.**	P18,20	Ver figura T15.	
T15	0	00	A lo ancho de la pantalla, abajo del planeta.			

Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T15	0	01	En el espacio alrededor del planeta, sólo abajo, no a lo ancho.			
T15	1	10	A lo ancho de la pantalla, arriba del planeta.			
T15	1	11	En el espacio alrededor del planeta, sólo arriba, no a lo ancho.			
T15	2	20	En la parte superior del planeta.			
T15	2	21	En la parte superior de algún/algunos continente/s.			
T15	3	30	En espacio alrededor del planeta o “por todas partes”.		Si T14 = 02: no dibuja cielo, seleccionar esta opción.	
T15	3	31	En el planeta.		Elegir esta opción si (no sólo si) representa el cielo con la misma circunferencia con la que representó el planeta.	
T15	3	32	Inmediatamente alrededor del planeta.		Inmediatamente a continuación del perímetro. No necesariamente rodeando todo el planeta.	
T16			**Ubicación de la Luna relativa al cielo.**	P12,14,18,20	Ver figura T20. Si T14 = 01 (dibuja nubes espontáneamente), traducir nubes a una línea según figura T20.	
T16	na	-01	Luna no disponible al concluir dibujo del cielo.			
T16	0	00	Dentro/más acá de, o superpuestas con, el cielo.		Elegir esta opción cuando T14 = 02 (no dibuja cielo) o T15 = 30 (cielo en espacio alrededor del planeta o “por todas partes”)	
T16	1	10	Fuera/más allá del cielo.			
T17			**Ubicación de las estrellas relativa al cielo.**	P13-16,18,20	Ver figura T20. Si T14 = 01 (dibuja nubes espontáneamente), traducir nubes a una línea según figura T20.	
T17	na	-01	Estrellas no disponibles al concluir dibujo del cielo.			
T17	0	00	Dentro/más acá de, o superpuestas con, el cielo.		Elegir esta opción cuando T14 = 02 (no dibuja cielo) o T15 = 30 (cielo en espacio alrededor del planeta o “por todas partes”)	
T17	0	01	Dentro o superpuestas, y fuera.		Intersección entre T17.00 y T17.10.	
T17	1	10	Fuera/más allá del cielo.			
T18			**Ubicación del Sol relativa al cielo.**	P17-18,20	Ver figura T20. Si T14 = 01 (dibuja nubes espontáneamente), traducir nubes a una línea según figura T20.	
T18	na	-01	Sol no disponible al concluir dibujo del cielo.			
T18	0	00	Dentro/más acá de, o superpuestas con, el cielo.		Elegir esta opción cuando T14 = 02 (no dibuja cielo) o T15 = 30 (cielo en espacio alrededor del planeta o “por todas partes”)	
T18	1	10	Fuera/más allá del cielo.			
T19			**Ubicación de las nubes.**	P19-22	Ver figura T19. Si más de una opción posible, seleccionar la menor.	
T19	0	00	A lo ancho de la pantalla, abajo del planeta.			
T19	0	01	En el espacio alrededor del planeta, sólo abajo, no a lo ancho.			

Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T19	1	10	A lo ancho de la pantalla, arriba del planeta.			
T19	1	11	En el espacio alrededor del planeta, sólo arriba, no a lo ancho.			
T19	2	20	En la parte superior del planeta.			
T19	2	21	En la parte superior de algún/algunos continente/s.			
T19	3	30	En espacio alrededor del planeta.			
T19	4	40	Distribuidas dentro del planeta.			
T19	4	41	Inmediatamente alrededor del planeta.		Inmediatamente a continuación del perímetro. No necesariamente rodeando todo el planeta.	
T20			**Ubicación de las nubes relativa al cielo.**	P18-P22	Ver figura T20.	
T20	0	00	Fuera/más allá del cielo.			“arriba del cielo”,
T20	0	01	Fuera, y dentro o superpuestas.		Intersección entre T20.00 y T20.02.	
T20	0	02	Dentro/más acá de, o superpuestas con, el cielo.		Elegir esta opción cuando T14 = 01 (dibuja nubes espontáneamente) o T14 = 02 (no dibuja cielo).	“abajo del cielo”, “en el cielo”,
T21			**Ubicación del sujeto.**	P24-25	Ver figura T21. Considerar dónde dibuja pies para evaluar “apoyado”. Considerar qué dibuja primero para evaluar “dentro/fuera”.	
T21	0	00	Fuera del planeta, apoyado en borde inferior de la pantalla.		Si la intención del dibujante es evidente, se considerará “apoyado” aún cuando no esté en contacto con el “suelo”.	
T21	0	01	Fuera del planeta, en el espacio alrededor de aquél.			
T21	1	10	Dentro del planeta, apoyado en la mitad inferior del perímetro o en un suelo interno.		Idem T21.00.	
T21	1	11	Dentro del planeta, dentro de un continente y apoyado en la mitad inferior no perimetral del borde o en un suelo interno.		Idem T21.00.	
T21	1	12	En la parte inferior del planeta, perpendicular o aproximadamente perpendicular al perímetro.		Totalmente dentro de la mitad inferior del planeta. Perpendicular al perímetro. Distancia al perímetro menor o igual que mitad de la altura (h), y libre (*) de otros objetos.	
T21	2	20	Dentro del planeta, libre o dentro de un continente.			
T21	2	21	Apoyado en el perímetro externo del planeta.			
T22			**Colocación continental del sujeto.**	P09-11,24-25	Entender “dentro de” como “sobre” en caso de un continente representado con una línea.	
T22	na	-01	Sin colocación continental.			
T22	0	00	Coloca sujeto dentro de un continente representado previamente.			
T22	0	01	Coloca sujeto dentro de un continente representado *ad hoc*.			
T23			**Ubicación de la casa.**	P26	Ver figura T23. Considerar dónde dibuja la base para evaluar “apoyada”. Considerar qué dibuja primero para evaluar “dentro/fuera”.	

Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T23	0	00	Fuera del planeta, apoyada en borde inferior de la pantalla.		Si la intención del dibujante es evidente, se considerará "apoyado" aún cuando no esté en contacto con el "suelo".	
T23	0	01	Fuera del planeta, en el espacio alrededor de aquél.			
T23	1	10	Dentro del planeta, apoyada en la mitad inferior del perímetro o en un suelo interno.		Idem T23.00.	
T23	1	11	Dentro del planeta, dentro de un continente y apoyado en la mitad inferior no perimetral del borde o en un suelo interno.		Idem T23.00.	
T23	1	12	En la parte inferior del planeta, perpendicular o aproximadamente perpendicular al perímetro.		Totalmente dentro de la mitad inferior del planeta. Perpendicular al perímetro. Distancia al perímetro menor o igual que mitad de la altura (h), y libre (*) de otros objetos.	
T23	2	20	Dentro del planeta, libre o dentro de un continente.			
T23	2	21	Apoyada en el perímetro externo del planeta.			
T24			**Ubicación del árbol.**	P27	Ver figura T24. Considerar dónde dibuja la base del tronco para evaluar "apoyado". Considerar qué dibuja primero para evaluar "dentro/fuera".	
T24	0	00	Fuera del planeta, apoyado en borde inferior de la pantalla.		Si la intención del dibujante es evidente, se considerará "apoyado" aún cuando no esté en contacto con el "suelo".	
T24	0	01	Fuera del planeta, en el espacio alrededor de aquél.			
T24	1	10	Dentro del planeta, apoyado en la mitad inferior del perímetro o en un suelo interno.		Idem T24.00.	
T24	1	11	Dentro del planeta, dentro de un continente y apoyado en la mitad inferior no perimetral del borde o en un suelo interno.		Idem T24.00.	
T24	1	12	En la parte inferior del planeta, perpendicular o aproximadamente perpendicular al perímetro.		Totalmente dentro de la mitad inferior del planeta. Perpendicular al perímetro. Distancia al perímetro menor o igual que mitad de la altura (h), y libre (*) de otros objetos.	
T24	2	20	Dentro del planeta, libre o dentro de un continente.			
T24	2	21	Apoyado en el perímetro externo del planeta.			
T25			**Representación del país en el que estamos.**	P09-11,29	Ver figura T25. No codificar como "punto" cuando con el punto señalado el sujeto refiere a un área que lo incluye.	
T25	0	00	Refiere a, o área apoyada en, el margen inferior de la pantalla.			
T25	0	01	Punto, línea o área fuera del planeta, en el espacio alrededor de aquél.			
T25	0	02	País y planeta son la misma cosa.			
T25	0	03	Fuera del planeta, dibuja contorno característico y/o menciona relaciones con otros países (limítrofes o no).			
T25	1	10	Punto o línea en el perímetro del planeta.			
T25	2	20	Punto o línea dentro del planeta.			
T25	2	21	Área perimetral dentro del planeta.		Entender "área perimetral" como aquella en la que el perímetro del planeta forma parte del perímetro del área.	
T25	2	22	Área no perimetral dentro del planeta.			

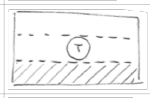
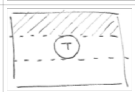
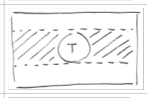
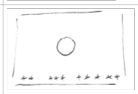
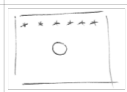
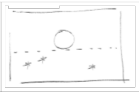
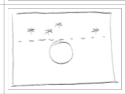
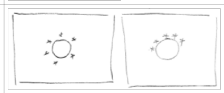
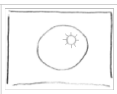
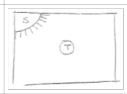
Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T25	3	30	Dentro del planeta, dibuja contorno característico y/o menciona relaciones con otros países (limítrofes o no).			
T26			**Inclusión de dibujos previos dentro del país.**	P24-27,29-31	En caso de haber representado país con una línea, entender "incluye dentro" como "apoya sobre".	
T26	0	00	No incluye sujeto, casa ni árbol dentro del país.		En caso de haber representado país con un un punto, seleccionar esta opción.	
T26	1	10	Incluye sólo árbol dentro del país.			
T26	2	20	Incluye sujeto y/o casa dentro del país.			
T26	2	21	Incluye sujeto, casa y árbol dentro del país.			
T27			**Ubicación de China relativa a Argentina.**	P29-32	Ver figura T27. Considerar líneas medias respecto a márgenes de la pantalla si dibujó país fuera del planeta en T25.	
T27	0	00	Sin cruzar ecuador ni meridiano.			
T27	1	10	Cruzando sólo ecuador.			
T27	2	20	Cruzando sólo meridiano.			
T27	2	21	Cruzando ecuador y meridiano.			
T27	3	30	Detrás de la cara visible.			
T28			**Ubicación de persona china.**	P33	Ver figura T21. Considerar dónde dibuja pies para evaluar "apoyado". Considerar qué dibuja primero para evaluar "dentro/fuera".	
T28	0	00	Fuera del planeta, apoyado en borde inferior de la pantalla.		Si la intención del dibujante es evidente, se considerará "apoyado" aún cuando no esté en contacto con el "suelo".	
T28	0	01	Fuera del planeta, en el espacio alrededor de aquél.			
T28	1	10	Dentro del planeta, apoyado en la mitad inferior del perímetro o en un suelo interno.		Idem T28.00.	
T28	1	11	Dentro del planeta, dentro de un continente y apoyado en la mitad inferior no perimetral del borde o en un suelo interno.		Idem T28.00.	
T28	1	12	En la parte inferior del planeta, perpendicular o aproximadamente perpendicular al perímetro.		Totalmente dentro de la mitad inferior del planeta. Perpendicular al perímetro. Distancia al perímetro menor o igual que mitad de la altura (h), y libre (*) de otros objetos.	
T28	2	20	Dentro del planeta, libre o dentro de un continente.			
T28	2	21	Apoyado en el perímetro externo del planeta.			
T29			**Visibilidad del Sol, la Luna y las estrellas.**	P34-35		
T29	0	00	No.			
T29	1	10	Parcialmente.			
T29	1	11	Sí.			

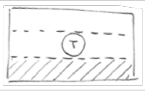
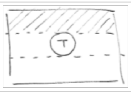
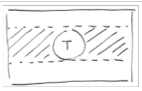
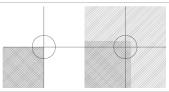
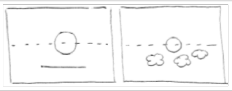
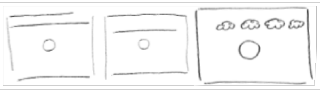
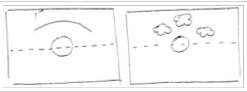
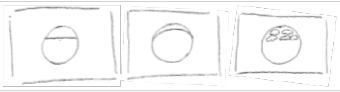
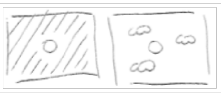
Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T30			**Aclaraciones sobre la visibilidad de los astros.**	P34-35		
T30	na	-01	Sin aclaraciones.			
T30	0	00	El Sol, la Luna y/o las estrellas están lejos y/o en el espacio.			"las estrellas están en el espacio", "puede verlos con un telescopio",
T30	0	01	Nosotros -observadores- estamos dentro del planeta.			"[ella] está en el planeta Tierra",
T30	0	02	Los astros se ven distinto (número, tamaño, forma, color, etc) según la ubicación del observador en el planeta.			"pueden ver la Luna, pero no el Sol", "el Sol está más cerca allí", "allí se ve más borroso",
T30	0	03	Allí -en China- los astros se ven a distinto tiempo que aquí.			
T30	0	04	Un observador sólo ve los astros de su lado del planeta.			
T31			**Ubicación del mar.**	P36	Ver figura T31.	
T31	0	00	Fuera del planeta, apoyado en borde inferior de la pantalla.			
T31	0	01	Fuera del planeta, en el espacio alrededor de aquél.			
T31	1	10	Apoyado en la mitad inferior del perímetro del planeta.			
T31	2	20	En el perímetro del planeta.			
T31	2	21	Dentro del planeta, área perimetral.		Entender "área perimetral" como aquella en la que el perímetro del planeta forma parte del perímetro del área.	
T31	3	30	Dentro del planeta, área no perimetral.			
T31	3	31	Dentro del planeta, entre China y Argentina.			
T31	4	40	Dentro del planeta, inundando todo el espacio no cubierto por continentes.			
T32			**Punto de inicio de la exploración.**	P38		
T32	0	00	Fuera del planeta.		Si el dibujo del planeta ha sido borrado, seleccionar esta opción.	
T32	1	10	Dentro del planeta.		Incluye puntos del perímetro.	
T33			**Interacción espontánea con el perímetro.**	P38-41		
T33	0	00	Repara espontáneamente en el perímetro.			
T33	0	01	No alcanza el perímetro.			
T33	1	10	Supera el perímetro sin reparar espontáneamente en él.		Incluye salidas al espacio y progresión por detrás del planeta sin interacción con el perímetro.	
T34			**Límites del espacio navegable.**	P38-43,45-47		

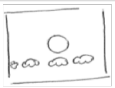
Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T34	0	00	Al menos un punto de la exploración se encuentra fuera del planeta.			
T34	0	01	Al menos un punto de la exploración fuera del planeta, a borde de una nave o cohete.			
T34	1	10	Permanece en el planeta o en el perímetro durante toda la exploración.			
T35			**Visibilidad del espacio navegable.**	P38-43,45-47	Comprende eventos que tienen al dibujo como marco de referencia.	
T35	0	00	Todos los puntos referidos durante la exploración se hallan en el plano y dentro de los límites del dibujo.			
T35	1	10	Algunos puntos referidos durante la exploración no se hallan en el plano o dentro de los límites del dibujo.			puntos en la otra cara del planeta, puntos más allá de los márgenes de la pantalla,
T36			**Desplazamientos desde el perímetro.**	P38-43,45-47	Si más de una opción posible, seleccionar la menor.	
T36	na	-01	Sin referencias al perímetro del planeta.		Usar sólo cuando T33 = 10 (“supera el perímetro sin reparar espontáneamente en él”), y P42 y P45 no refieren al perímetro del planeta.	
T36	0	00	Partiendo del perímetro, abandona el planeta.			desde el perímetro del planeta sale al espacio,
T36	1	10	Partiendo del perímetro, sólo se detiene en, o se desplaza por, aquél.			
T36	2	20	Partiendo del perímetro, reaparece en un punto perimetral opuesto.			
T36	2	21	Partiendo del perímetro, permanece en el planeta.		Incluye progresión hacia atrás del planeta desde el perímetro.	
T37			**Perímetro del planeta como límite.**	P38-43,45-47		
T37	na	-01	Sin referencias al perímetro del planeta.		Usar sólo cuando T33 = 10 (“supera el perímetro sin reparar espontáneamente en él”), y P42 y P45 no refieren al perímetro del planeta.	
T37	0	00	El perímetro es un límite.			
T37	1	10	El perímetro es un punto ordinario en el planeta.			un lugar en el mar, un punto en un país,
T38			**Naturaleza del límite perimetral.**	P38-48	Codificar como el producto de los valores de todas las opciones válidas.	
T38	na	-01	Sin referencia al perímetro como límite.		Seleccionar si T37 = 10 (el perímetro es un punto ordinario en el planeta).	
T38	0	02	El perímetro es la orilla del mar.			
T38	0	03	El perímetro es el cielo.			
T38	0	05	El perímetro es un límite cerrado, una barrera.			
T38	0	07	El perímetro es el espacio.			
T38	0	11	El perímetro es un final.			“no hay más camino”,
T38	0	13	El perímetro es un abismo.			es posible caer fuera del planeta desde el perímetro,

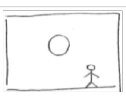
Tema	Niv	Cód	Descripción	Preguntas	Figuras/Indicaciones	Ejemplos
T39			**Finitud del planeta.**	P38-48		
T39	0	00	El planeta se termina en los márgenes de la pantalla.			
T39	1	10	El planeta se termina.			
T39	1	11	El planeta se termina en el perímetro.			
T39	2	20	El planeta no se termina.		Seleccionar sólo si en ningún momento expresó la existencia de un borde o final del planeta.	
T39	2	21	El planeta no se termina porque no es plano o es esférico.			
T40			**Ciclo día/noche.**	-		
T40	na	-01	Sin mención.			
T40	0	00	El día y la noche dependen de la posición en el planeta.			"la noche está acá", "cuando en China es de día, en Argentina es de noche", "el Sol ilumina sólo una parte del planeta",
T41			**Limitaciones de escala.**	-		
T41	na	-01	Sin mención.			
T41	0	00	Manifiesta limitaciones de escala o de espacio en el dibujo.			"la Luna está más lejos que donde la dibujé", "el Sol es más grande que como lo dibujé",
T42			**Refiere al espacio oculto detrás del planeta.**	-		
T42	na	-01	Sin mención.			
T42	0	00	Refiere al espacio o a objetos ocultos total o parcialmente detrás de la cara visible del planeta.			dibuja nubes parcialmente ocultas detrás del planeta, menciona países detrás de la cara visible del planeta,
T43			**Arriba y abajo en el planeta.**	P37		
T43	na	-01	Sin mención.			
T43	0	00	Las cosas caen desde regiones superiores hacia regiones inferiores del planeta.			
T43	0	01	Expresa "hacia regiones superiores del planeta" como "hacia arriba", pero en sentido figurado, en referencia al dibujo.			
T43	0	02	Expresa que regiones claramente superiores del planeta no están encima de ni caen hacia regiones inferiores.			
T44			**Marco de la pantalla.**	-		
T44	na	-01	Sin mención.			
T44	0	00	El marco de la pantalla representa un límite real en el espacio.			
T44	1	10	El marco de la pantalla no representa un límite real en el espacio.			

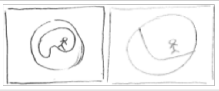
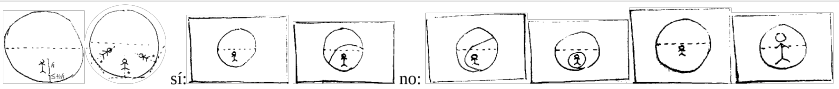
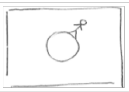
El código "-01" se reserva para *temas* sin codificar.

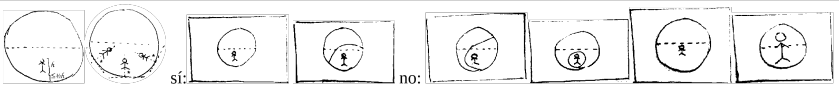
T07	**Ubicación de la Luna.**	
T07.00	Dentro del planeta.	
T07.10	Abajo del planeta.	
T07.11	Arriba del planeta.	
T07.12	A los lados del planeta.	
T07.13	Girando alrededor del planeta.	
T09	**Ubicación de las estrellas.**	
T09.00	Dentro del planeta.	
T09.10	A lo ancho de la pantalla, abajo del planeta.	
T09.11	A lo ancho de la pantalla, arriba del planeta.	
T09.20	Sólo abajo del planeta, no a lo ancho.	
T09.21	Sólo arriba del planeta, no a lo ancho.	
T09.22	Sólo a los lados del planeta.	
T09.30	Inmediatamente alrededor del planeta.	
T09.31	En espacio alrededor del planeta o por todas partes.	
T10	**Asociación entre estrellas y Luna y/o noche.**	
T10.00	Asocia la ubicación de las estrellas con la posición de la Luna y/o de la noche.	
T11	**Ubicación del Sol.**	
T11.00	Dentro del planeta.	
T11.10	Cuarto de Sol en esquina superior izquierda o derecha.	

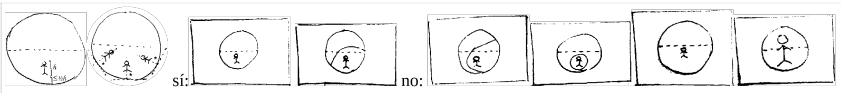
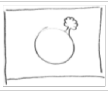
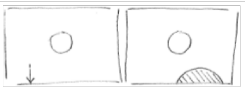
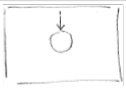
T11.20	Girando alrededor del planeta.	
T11.30	Abajo del planeta.	
T11.31	Arriba del planeta.	
T11.32	A los lados o atrás del planeta.	
T11.33	En el centro de la órbita terrestre.	
T13	**Ubicación del Sol relativa a la Luna.**	
T13.00	Sin cruzar líneas medias.	
T13.01	Cruzando sólo línea media horizontal.	
T13.02	Cruzando sólo línea media vertical.	
T13.03	Cruzando líneas medias vertical y horizontal.	
T15	**Ubicación del cielo.**	
T15.00	A lo ancho de la pantalla, abajo del planeta.	
T15.01	En el espacio alrededor del planeta, sólo abajo, no a lo ancho.	
T15.10	A lo ancho de la pantalla, arriba del planeta.	
T15.11	En el espacio alrededor del planeta, sólo arriba, no a lo ancho.	
T15.20	En la parte superior del planeta.	
T15.21	En la parte superior de algún/algunos continente/s.	
T15.30	En espacio alrededor del planeta o “por todas partes”.	

T15.31	En el planeta.	
T15.32	Inmediatamente alrededor del planeta.	
T19	**Ubicación de las nubes.**	
T19.00	A lo ancho de la pantalla, abajo del planeta.	
T19.01	En el espacio alrededor del planeta, sólo abajo, no a lo ancho.	
T19.10	A lo ancho de la pantalla, arriba del planeta.	
T19.11	En el espacio alrededor del planeta, sólo arriba, no a lo ancho.	
T19.20	En la parte superior del planeta.	
T19.21	En la parte superior de algún/algunos continente/s.	

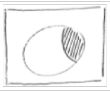
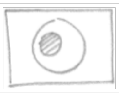
T19.30	En espacio alrededor del planeta.	
T19.40	Distribuidas dentro del planeta.	
T19.41	Inmediatamente alrededor del planeta.	
T20	**Ubicación de las nubes relativa al cielo.**	
T20.00	Fuera/más allá del cielo.	
T20.01	Fuera, y dentro o superpuestas.	
T20.02	Dentro/más acá de, o superpuestas con, el cielo.	
T21	**Ubicación del sujeto.**	
T21.00	Fuera del planeta, apoyado en borde inferior de la pantalla.	
T21.01	Fuera del planeta, en el espacio alrededor de aquél.	

T21.10	Dentro del planeta, apoyado en la mitad inferior del perímetro o en un suelo interno.	
T21.11	Dentro del planeta, dentro de un continente y apoyado en la mitad inferior no perimetral del borde o en un suelo interno.	
T21.12	En la parte inferior del planeta, perpendicular o aproximadamente perpendicular al perímetro.	 sí: no:
T21.20	Dentro del planeta, libre o dentro de un continente.	
T21.21	Apoyado en el perímetro externo del planeta.	
T23	**Ubicación de la casa.**	
T23.00	Fuera del planeta, apoyada en borde inferior de la pantalla.	
T23.01	Fuera del planeta, en el espacio alrededor de aquél.	
T23.10	Dentro del planeta, apoyada en la mitad inferior del perímetro o en un suelo interno.	

T23.11	Dentro del planeta, dentro de un continente y apoyado en la mitad inferior no perimetral del borde o en un suelo interno.	
T23.12	En la parte inferior del planeta, perpendicular o aproximadamente perpendicular al perímetro.	 sí: no:
T23.20	Dentro del planeta, libre o dentro de un continente.	
T23.21	Apoyada en el perímetro externo del planeta.	
T24	**Ubicación del árbol.**	
T24.00	Fuera del planeta, apoyado en borde inferior de la pantalla.	
T24.01	Fuera del planeta, en el espacio alrededor de aquél.	
T24.10	Dentro del planeta, apoyado en la mitad inferior del perímetro o en un suelo interno.	
T24.11	Dentro del planeta, dentro de un continente y apoyado en la mitad inferior no perimetral del borde o en un suelo interno.	

T24.12	En la parte inferior del planeta, perpendicular o aproximadamente perpendicular al perímetro.	
T24.20	Dentro del planeta, libre o dentro de un continente.	
T24.21	Apoyado en el perímetro externo del planeta.	
T25	**Representación del país en el que estamos.**	
T25.00	Refiere a, o área apoyada en, el margen inferior de la pantalla.	
T25.01	Punto, línea o área fuera del planeta, en el espacio alrededor de aquél.	
T25.02	País y planeta son la misma cosa.	
T25.03	Fuera del planeta, dibuja contorno característico y/o menciona relaciones con otros países.	
T25.10	Punto o línea en el perímetro del planeta.	

T25.20	Punto o línea dentro del planeta.	
T25.21	Área perimetral dentro del planeta.	
T25.22	Área no perimetral dentro del planeta.	
T25.30	Dentro del planeta, dibuja contorno característico y/o menciona relaciones con otros países.	
T27	**Ubicación de China relativa a Argentina.**	
T27.00	Sin cruzar ecuador ni meridiano.	
T27.10	Cruzando sólo ecuador.	
T27.20	Cruzando sólo meridiano.	
T27.21	Cruzando ecuador y meridiano.	
T31	**Ubicación del mar.**	

T31.00	Fuera del planeta, apoyado en borde inferior de la pantalla.	
T31.01	Fuera del planeta, en el espacio alrededor de aquél.	
T31.10	Apoyado en la mitad inferior del perímetro del planeta.	
T31.20	En el perímetro del planeta.	
T31.21	Dentro del planeta, área perimetral.	
T31.30	Dentro del planeta, área no perimetral.	
T31.31	Dentro del planeta, entre China y Argentina.	
T31.40	Dentro del planeta, inundando todo el espacio no cubierto por continentes.	

Anexo 2.3: Correspondencia entre preguntas y temas

Correspondencia de las 50 preguntas (P) del cuestionario con los 44 *temas* (T) del Esquema de codificación 2. Las cruces (x) indican preguntas cuyas respuestas fueron usadas para codificar los *temas* indicados. Las preguntas sin cruces no se usaron para codificar ningún *tema* específico. Los *temas* sin cruces refieren a cuestiones aludidas espontáneamente por les participantes en cualquier momento del cuestionario.

	T01	T02	T03	T04	T05	T06	T07	T08	T09	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20	T21	T22	T23	T24	T25	T26	T27	T28	T29	T30	T31	T32	T33	T34	T35	T36	T37	T38	T39	T40	T41	T42	T43	T44
P01																																											
P02																																												
P03																																												
P04																																												
P05	X																																											
P06		X	X		X																																						
P07		X	X		X																																						
P08		X	X		X																																						
P09				X	X																	X			X																		
P10				X	X																	X			X																		
P11				X	X																	X			X																		
P12						X	X	X		X			X			X																												
P13									X	X							X																											
P14							X	X	X	X			X			X	X																												
P15									X	X							X																												
P16									X	X							X																												
P17											X	X	X					X																											
P18														X	X	X	X	X		X																									
P19																			X	X																								
P20											X	X	X	X	X	X	X	X	X	X																								
P21																			X	X																								
P22																			X	X																								
P23																																												
P24																					X	X				X																		
P25																					X	X				X																		
P26																							X			X																		
P27																								X		X																		
P28																												X																
P29																									X	X	X																		
P30																										X	X																		
P31																										X	X																		
P32																											X																		
P33																												X																	
P34																													X	X															
P35																													X	X															
P36																															X														
P37																																												X	
P38																																X	X	X	X	X	X	X	X						
P39																																	X	X	X	X	X	X	X						
P40																																	X	X	X	X	X	X	X						
P41																																	X	X	X	X	X	X	X						
P42																																		X	X	X	X	X	X						
P43																																		X	X	X	X	X	X						
P44																																						X	X						
P45																																		X	X	X	X	X	X						
P46																																		X	X	X	X	X	X						
P47																																		X	X	X	X	X	X						
P48																																						X	X						
P49																																													
P50																																													

Anexo 2.4: Temas del Esquema de codificación 2

Las respuestas que dieron les participantes a las preguntas del cuestionario en la Entrevista 2 fueron reunidas y usadas para definir los 44 *temas* (T) de conocimiento enumerados a continuación.

Núm	Descripción	Tipo (1)	Códs. (2)	Nivs. (3)	Punt. (4)	Ambas codif. (5)	Dimensión (6)	Dominio (7)
T01	Hacia dónde mirar para ver el planeta Tierra.	obl	4	4	sí	34	ubic_tierra	Cielo
T02	Forma abstracta del planeta.	obl	4	3	sí	42	forma_tierra	Cielo
T03	Forma análoga al planeta.	obl	4	3	sí	36	forma_tierra	Cielo
T04	Dibujo del planeta como se vería desde el espacio.	obl	3	2	sí	42	forma_tierra	Cielo
T05	Contradicción entre planeta redondo y plano.	opc	2	2	sí	1	forma_tierra	Cielo
T06	Forma de la Luna.	obl	4	3	sí	41	forma_luna	Cielo
T07	Ubicación de la Luna.	obl	5	2	sí	41	ubic_astro	Cielo
T08	Múltiples posiciones de la Luna.	opc	1	1	no	-	-	-
T09	Ubicación de las estrellas.	obl	8	4	sí	42	ubic_astro	Cielo
T10	Asociación entre estrellas y Luna y/o noche.	opc	1	1	no	-	-	-
T11	Ubicación del Sol.	obl	7	4	sí	42	ubic_astro	Cielo
T12	Múltiples posiciones del Sol.	opc	1	1	no	-	-	-
T13	Ubicación del Sol relativa a la Luna.	obl	4	1	no	-	-	-
T14	Representación del cielo.	obl	4	1	no	-	-	-
T15	Ubicación del cielo.	obl	9	4	sí	41	cielo	Cielo
T16	Ubicación de la Luna relativa al cielo.	opc	2	2	sí	39	cielo	Cielo
T17	Ubicación de las estrellas relativa al cielo.	opc	3	2	sí	41	cielo	Cielo
T18	Ubicación del Sol relativa al cielo.	opc	2	2	sí	40	cielo	Cielo
T19	Ubicación de las nubes.	obl	9	5	sí	42	cielo	Cielo
T20	Ubicación de las nubes relativa al cielo.	obl	3	1	no	-	-	-
T21	Ubicación del sujeto.	obl	7	3	sí	41	ubic_auto	Suelo
T22	Colocación continental del sujeto.	opc	2	1	no	-	-	-
T23	Ubicación de la casa.	obl	7	3	sí	32	ubic_auto	Suelo
T24	Ubicación del árbol.	obl	7	3	sí	36	ubic_auto	Suelo
T25	Representación del país en el que estamos.	obl	9	4	sí	42	país	Suelo
T26	Inclusión de dibujos previos dentro del país.	obl	4	3	sí	40	país	Suelo
T27	Ubicación de China relativa a Argentina. (*)	obl	5	4	sí	42	china	Suelo
T28	Ubicación de persona china.	obl	7	3	sí	42	ubic_auto	Suelo
T29	Visibilidad del Sol, la Luna y las estrellas, desde China.	obl	3	2	sí	37	china	Cielo
T30	Aclaraciones sobre la visibilidad de los astros.	opc	5	1	no	-	-	-
T31	Ubicación del mar.	obl	8	5	sí	40	mar	Suelo
T32	Punto de inicio de la exploración.	obl	2	2	sí	42	exploración	Suelo
T33	Interacción espontánea con el perímetro del planeta.	obl	3	2	sí	42	bordes	Exploración
T34	Límites del espacio navegable.	obl	3	2	sí	42	exploración	Exploración
T35	Visibilidad del espacio navegable.	obl	2	2	sí	42	exploración	Exploración
T36	Desplazamientos desde el perímetro.	opc	4	3	sí	13	bordes	Exploración
T37	Perímetro del planeta como límite.	opc	2	2	sí	21	bordes	Exploración
T38	Naturaleza del límite perimetral.	opc	6	1	no	-	-	-
T39	Finitud del planeta.	obl	5	3	sí	36	bordes	Exploración
T40	Ciclo día/noche.	opc	1	1	no	-	-	-
T41	Limitaciones de escala.	opc	1	1	no	-	-	-
T42	Refiere al espacio oculto detrás del planeta.	opc	1	1	no	-	-	-
T43	Arriba y abajo en el dibujo del planeta.	opc	3	1	no	-	-	-
T44	Significado del marco de la pantalla.	opc	2	2	sí	4	exploración	Exploración

(1) Tipo: Los temas que corresponden a aspectos explícitamente referidos en el cuestionario eran *de codificación obligada* (obl). Los que refieren a cuestiones que podían o no ser aludidas espontáneamente por les participantes eran *de codificación opcional* (opc).

(2) Códigos: Número de códigos predefinidos para un tema dado. Los *temas* con sólo un código predefinido eran o bien codificados o bien no codificados.

(3) Niveles. Número de niveles de conocimiento en los que los códigos podían ser ordenados.

(4) Puntuado: Los *temas* con códigos predefinidos en dos o más niveles (31 *temas* en total) fueron puntuados (*sí*). Dentro de cada *tema* los códigos correspondientes a respuestas más avanzadas recibieron un puntaje más alto. Códigos en un mismo nivel de conocimiento recibieron el mismo puntaje.

(5) Ambas codificadas: Número de participantes para les que cada *tema* fue codificado en ambas entrevistas. Sólo los *temas* codificados en ambas entrevistas para al menos 40 participantes fueron considerados en las diferencias de puntaje.

(6) Dimensión: Los *temas* puntuados fueron combinados en 11 dimensiones. Se calcularon tres puntajes promedio por dimensión (uno por cada canal), y luego fueron promediados nuevamente para obtener un puntaje promedio por dimensión.

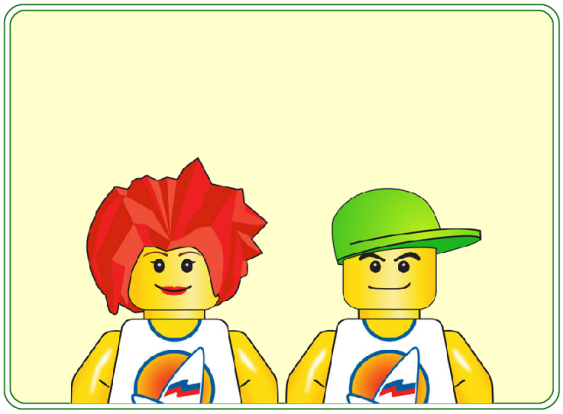
(7) Dominios: Alternativamente, los *temas* puntuados fueron agrupados en tres *dominios* (ver “2.2.3.4 Efecto en los modelos mentales de la Tierra”).

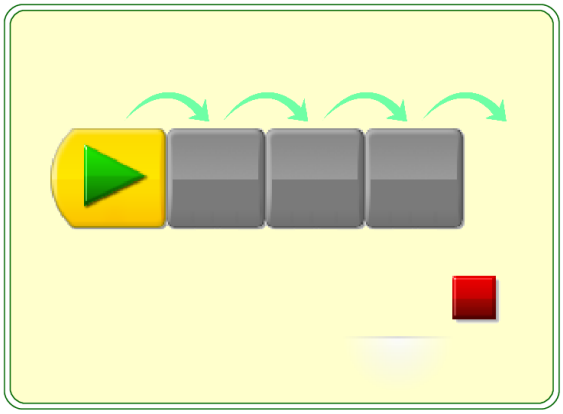
(*) Se usó China porque su cualidad de antípoda respecto de Argentina (donde se realizó el estudio) está ampliamente presente en la cultura popular.

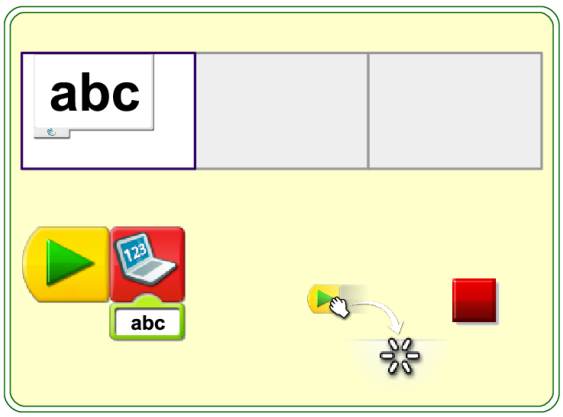
Dimensiones: *ubic_tierra*: ubicación de la Tierra; *forma_tierra*: forma de la Tierra; *forma_luna*: forma de la luna; *ubic_astro*: ubicación de otros cuerpos celestes; *cielo*: ubicación del cielo; *ubic_auto*: ubicación de le participante en la Tierra; *país*: representación y ubicación del país; *ubic_china*: ubicación de la antípoda; *mar*: ubicación del mar; *exploración*: exploración de la superficie terrestre; *bordes*: bordes del planeta.

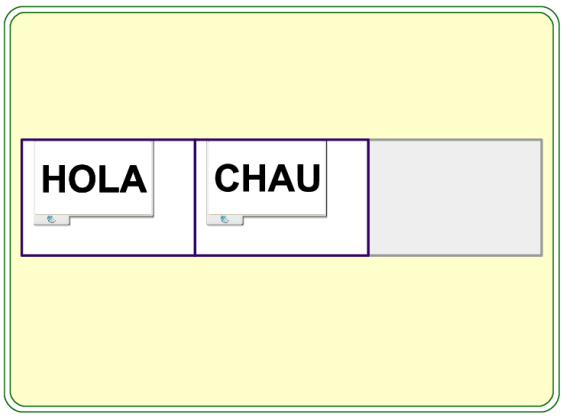
Anexo 3.1: Catálogo de imágenes y sonidos

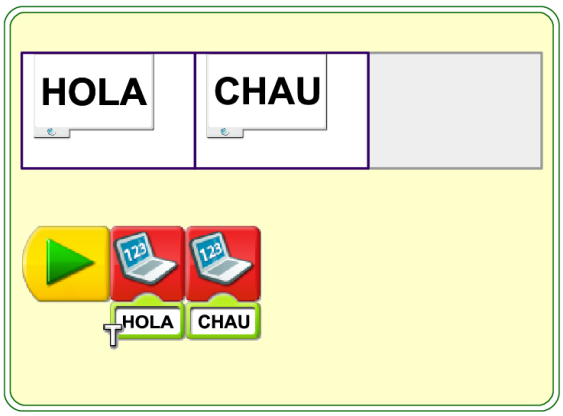
Anexo 3.2: Libro de instrucciones

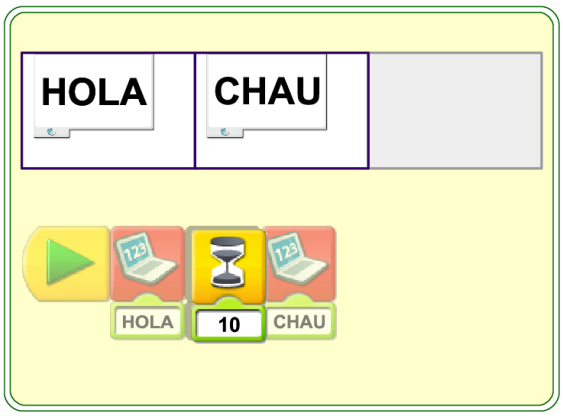


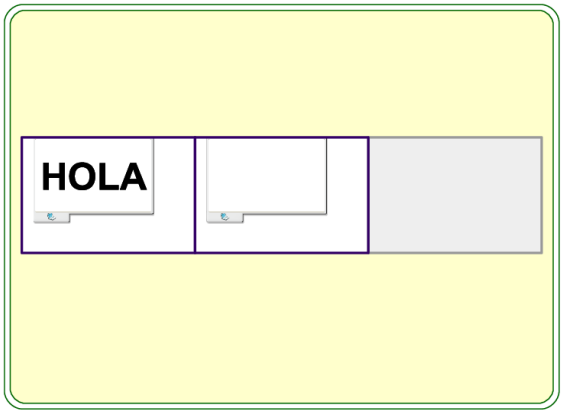




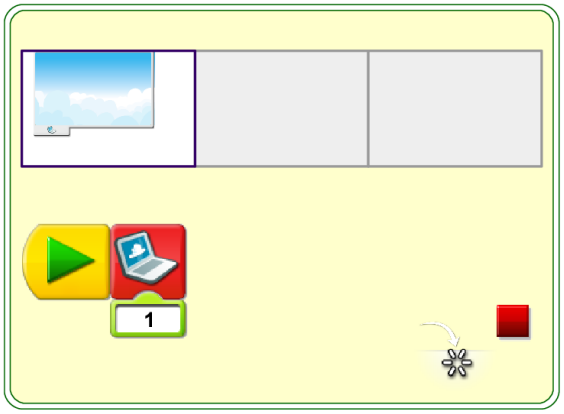




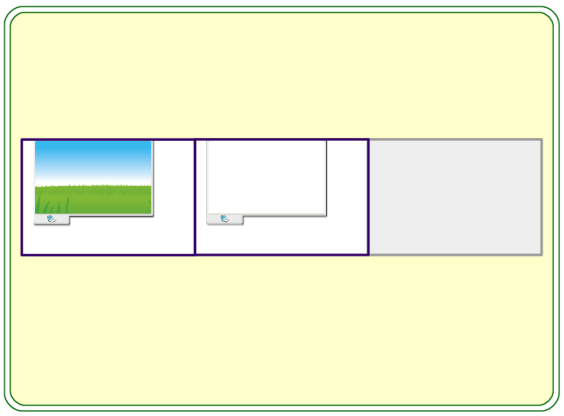


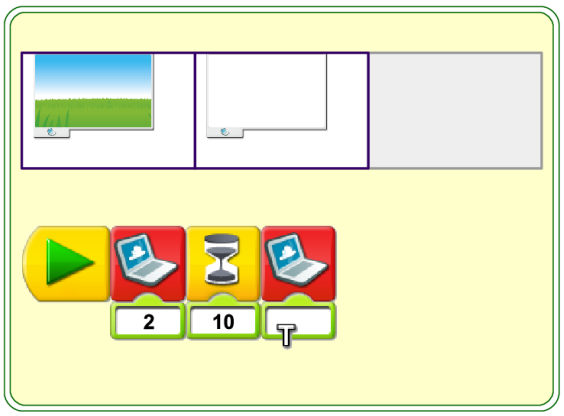

HOLA

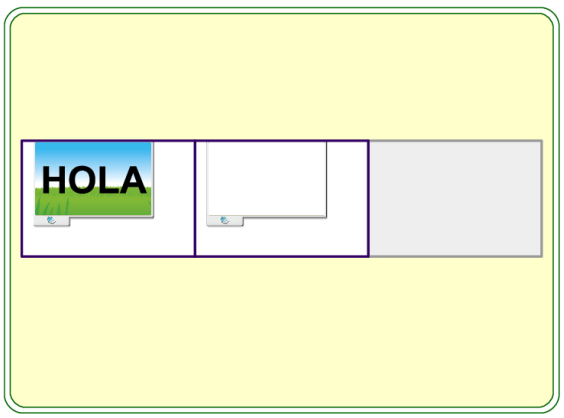
HOLA
HOLA
10


1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

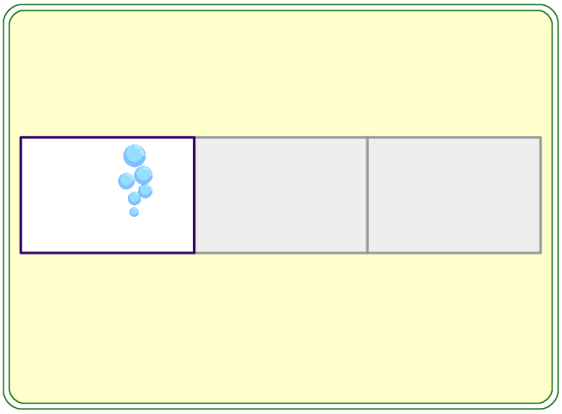


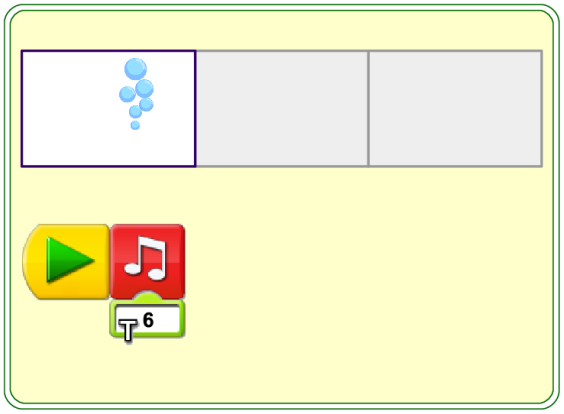

2
10

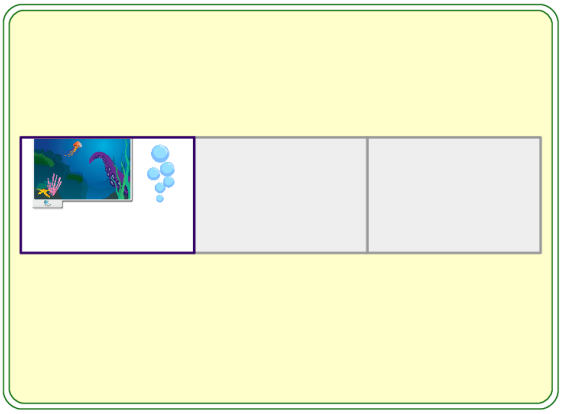

HOLA

HOLA
2
HOLA
10

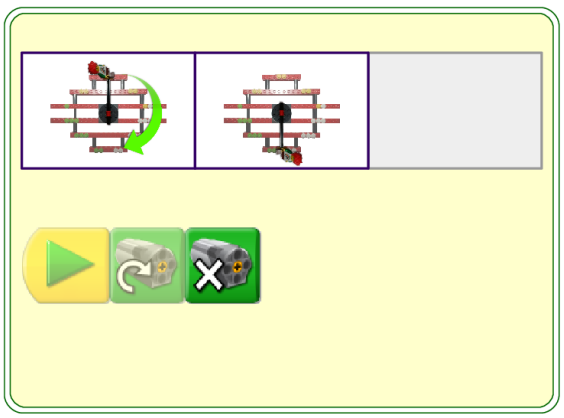
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

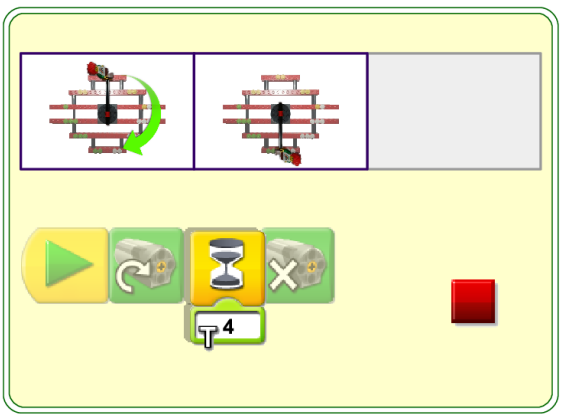


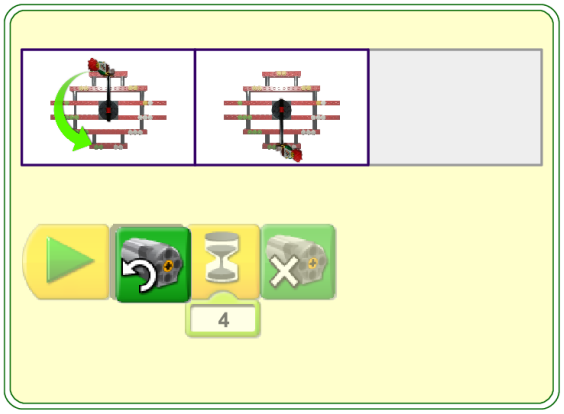

6



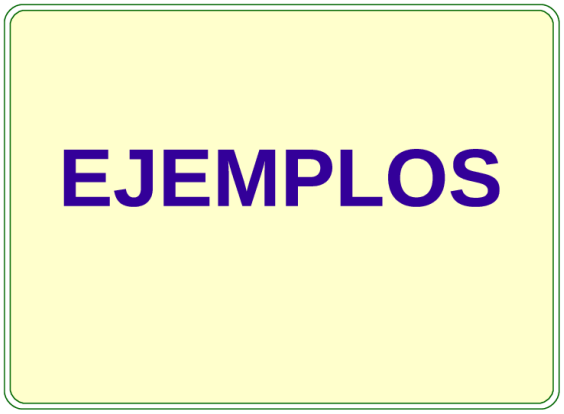
8
6




4


4

Anexo 3.3: Libro de ejemplos



Anexo 3.4: Eventos en videos

En cada evento se muestran sólo las variables que cambian.

El estado inicial es el mismo para todos los ejercicios: [img=0, txt="", pos=0, rot=0, snd=0].

La secuencia correcta se muestra según la tabla de traducción (ver “Anexo 3.8: Tabla de traducción distancias de Levenshtein”)

	Eventos o cambios de estado (evaluación/reevaluación)				Secuencia correcta (evaluación / reevaluación)
	Evento 1	Evento 2	Evento 3	Evento 4	
	1seg	3seg	5seg	7seg	
Ejercicio 1	txt=HOLA/SALUT	txt=""			STlWwT_ / STvWwT_
Ejercicio 2	img=5 / 7	img=11 / 10	img=0		SB5WwBbWwB_ / SB7WwBaWwB_
Ejercicio 3	snd=9 / 15	snd=0 txt=PATO / RING	txt=""		SP9TmWwT_ / SPfTxWwT_
Ejercicio 4	img=14 / 18 snd=2 / 20	snd=0 txt=RANA / AUTO	img=0 txt=""		SBeP2TnWwB_T_ / SBlPkTyWwB_T_
Ejercicio 5	txt=HEY / UY snd=12 / 7	snd=0 rot=+1 / -1	rot=0 pos=180 txt=""		SToPcRWwHT_ / STzP7LWwHT_
Ejercicio 6	snd=14 / 19	snd=0 txt=LEÓN / AVES	txt="" img=6 / 12 snd=14 / 19	snd=0 img=0	SPeTpWwT_B6PeB_ / SPjTAWwT_BcPjB_
Ejercicio 7	img=15 / 13 txt=321 / LA	txt=YA / CLASE rot=-1 / +1	rot=0 pos=0 img=0 txt=""		SBfTqWwTrLW9HB_T_ / SBdTCWwTDRWwHB_T_
Ejercicio 8	txt=EL / UN	txt=LICEO / OSO	img=16 / 9 snd=11 / 13	img=0 txt="" snd=0	STsWwTtWwBgPbB_T_ / STEWwTFWwB9PdB_T_
Ejercicio 9	rot=+1 / -1	rot=0 pos=180 img=17 / 5 snd=5 / 4	snd=0 rot=-1 / +1	rot=0 pos=0 img=0	SRWwHBhP5LWwHB_ / SLW4HB5P4RW4HB_
Ejercicio 10	img=19 / 3	img=20 / 4 txt=ACHÚ / BRUM snd=18 / 10	snd=0 txt="" rot=+1 / -1	rot=0 pos=0 img=0	SBjWwBkTuPlT_RWwHB_ / SB3WwB4TGPaT_LWwHB_

Anexo 3.5: Guión de instrucción

Discurso empleado en la etapa de Instrucción (ver apartado Etapa de Instrucción *de la sección "3.2.1.5 Procedimiento").*

Las indicaciones no verbales, y los números de página de los libros de instrucciones (Parte 1; ver "Anexo 3.2: Libro de instrucciones") y de ejemplos (Parte 2; ver "Anexo 3.3: Libro de ejemplos"), se indican entre corchetes.

Siempre que les participantes señalaban un bloque o un lugar en la pantalla, éste era señalado nuevamente por le instructore adulte para el resto de les participantes.

Parte 1

Les voy a enseñar a jugar un juego.

En el juego, hay que usar estos bloques [señalar panel inferior de la pantalla] para decirle a la compu lo que queremos que haga. Cada uno de estos bloques le dice algo distinto a la compu.

[1]

Este bloque [señalar en cuadernillo] es el bloque de inicio. Le dice a la compu dónde tiene que empezar a leer.

Entonces, la compu va a empezar acá, y después va a mirar el bloque que esté justo al lado. Después va a seguir con éste, después con éste, después con éste, y cuando ya no tenga más bloques va a terminar.

En cualquier momento, podemos usar el botón rojo [señalar cuadernillo y pantalla] para pedirle a la compu que arranque otra vez de cero.

[2]

¿Qué letras son éstas [señalar]? "a", "b", "c".

¿Cómo hago para que la compu escriba acá [señalar en pantalla] "abc"?

Primero, pongo el bloque de inicio [señalar en cuadernillo].

Después, este bloque [señalar en cuadernillo] le dice a la compu lo que tiene que escribir.

Entonces, a ver, ¿dónde está el bloque de inicio en la pantalla? [señalan] Lo agarro y lo pongo acá [soltar en lienzo del editor].

¿Y ahora? ¿Qué bloque tengo que poner? (¿Dónde está? ¿Dónde está el bloque que tengo que poner para decirle a la compu que escriba?) [señalan] Muy bien. Lo agarro y lo engancho acá, justo al lado.

¿Ya está? Ahora hago clic en el bloque de inicio para pedirle a la compu que haga todo lo que le pedimos. ¿Probamos? [probar secuencia]

¿Funcionó? ¿Escribió “abc” acá [señalar en pantalla]? ¡Buenísimo!

Para borrar el programa que hicimos [señalar en cuaderno], lo agarro del bloque de inicio [agarrar], y lo llevo acá abajo [borrar].

Aprieto el botón rojo [señalar en cuaderno] para que la compu arranque otra vez de cero.

[3]

¿Cómo hago para que primero escriba “HOLA” [señalar viñeta 1], y después escriba “CHAU” [señalar viñeta 2]?

¿Cuál es el bloque que siempre ponemos primero? (¿Qué bloque le dice a la compu dónde empezar?) El bloque de inicio. Lo agarro y lo pongo acá.

Ahora, ¿qué bloque pongo para que escriba “HOLA? (¿cuál pusimos antes para que escriba “abc”?) Éste [señalar]. Lo agarro y lo engancho acá al lado.

Pero, si lo dejo así, ¿qué va a escribir? [señalar texto “abc” por defecto]. Entonces, ¿qué tengo que poner acá? [señalar bloque de argumento] Voy con la flechita acá -¿ven que se convierte en una letra?- y escribo “HOLA”.

¿Y después, qué tiene que pasar [señalar viñeta 2]? [responden] Tiene que escribir “CHAU”. Agarro otra vez el mismo bloque, lo engancho al lado, ¿y acá qué pongo? [responden] Escribo “CHAU”.

A ver…

[4]

¡Muy bien! ¿Lo probamos? Para probarlo, hago clic en el bloque de inicio. [pulsar bloque de inicio]

¿Qué pasó? ¿Funcionó? ¿Escribió “CHAU”? Sí, ¿pero escribió “HOLA”? A ver, probemos de nuevo. Prestemos atención acá [señalar pantalla; pulsar bloque de inicio nuevamente]

¿Escribió “HOLA”? ¡Fue muy rápido! ¿Qué tenemos que hacer antes de escribir “CHAU”? Tenemos que esperar.

[5]

Para decirle a la compu que espere un poquito, usamos este bloque. Entonces, lo agarramos y lo acomodamos acá [entre bloques], antes de decirle a la compu que escriba “CHAU”.

¿Probamos de vuelta? A ver… [pulsar bloque de inicio]

¿Y ahora funcionó? ¿Escribió “HOLA” [señalar viñeta]? ¡Sí! ¿Escribió “CHAU” [señalar viñeta]? ¡Sí! ¡Muy bien!

[6]

¿Cómo hacemos ahora para que después de escribir “HOLA” [señalar viñeta] escriba en blanco [señalar viñeta]?

[7]

Para eso, usamos el mismo bloque; y para que escriba en blanco borramos lo que dice acá abajo [señalar bloque de argumento]. Entonces, voy con la flechita hasta acá y borro con la tecla “BORRAR”.

¿Probamos? [pulsar bloque inicio] ¿Escribió “HOLA” primero [señalar viñeta]? ¿Escribió en blanco después [señalar viñeta]? ¡Muy bien!

Para borrar el programa que hicimos [señalar en cuaderno], lo agarro [agarrar el programa] y lo llevo acá abajo [borrar].

[8]

¿Como hacemos ahora para que la compu muestre este fondo [señalar viñeta]?

Este bloque [señalar en cuaderno] le dice a la compu que fondo tiene que mostrar.

Entonces, ¿qué bloque ponemos primero siempre? El bloque de inicio, muy bien. Lo agarro y lo pongo acá.

Ahora, ¿dónde está (en la pantalla) el bloque que le dice a la compu que fondo tiene que mostrar? Lo agarro y lo engancho al lado.

¿Probamos? Hago clic en el bloque de inicio. ¿Funcionó? ¿Mostró el fondo que queríamos? ¡Muy bien!

Antes de seguir, borramos lo que hicimos [señalar en cuaderno]: llevo el programa acá abajo [borrar]; y aprieto el botón rojo [señalar en cuaderno] para que la compu arranque otra vez de cero [pulsar rojo].

¿Hay más fondos…?

[9]

¡Sí! ¡Un montón de fondos más! [señalar en cuaderno] Y cada fondo tiene su número, ¿ven? [señalar dos primeras filas de números]

[10]

¿Cómo hacemos para que primero muestre el fondo del pasto [señalar viñeta], y después muestre el fondo vacío [señalar viñeta]?

¿Qué bloque ponemos primero siempre? El bloque de inicio, muy bien. [colocar]

Ahora, ¿qué bloque ponemos para que muestre un fondo? Lo agarro y lo engancho al lado.

Pero, ¿este fondo [señalar en cuaderno] es el fondo número 1 [señalar bloque de argumento]? ¿Cuál es el número del fondo del pasto? ¿Nos fijamos? [volver al catálogo de fondos] El fondo del pasto es el número 2 [señalar en cuaderno]

Entonces, ¿qué número tenemos que poner acá [señalar bloque de argumento]? Voy con la flechita y escribo “2”.

¿Y después qué tiene que pasar [señalar viñeta]? Tiene que mostrar el fondo vacío. Para eso, uso otra vez el mismo bloque [enganchar al lado], y para que muestre el fondo vacío voy con la flechita acá y borro el número que está acá abajo con la tecla “BORRAR”.

¿Ya está? ¿Probamos? [si alguien entiende que falta bloque de espera, probar de todos modos]

¿Qué pasó? ¿Funcionó? ¿Mostró el fondo vacío? Sí, ¿pero mostró el fondo del pasto? A ver, probemos de nuevo. Prestemos atención acá [señalar pantalla; pulsar bloque de inicio nuevamente]

¿Mostró el fondo del pasto? ¡Fue muy rápido otra vez! Entonces, ¿qué tenemos que hacer antes de mostrar el fondo vacío? Muy bien, tenemos que esperar.

¿Cuál era el bloque que le dice a la compu que espere un poquito? Lo agarro y lo acomodo acá, antes de decirle a la compu que muestre el fondo vacío.

A ver…

[11]

¿Probamos de nuevo? [pulsar bloque de inicio]

¿Funcionó? ¿Mostró el fondo del pasto primero [señalar viñeta]? ¿Mostró el fondo vacío después [señalar viñeta]? ¡Muy bien!

[12]

Ahora, ¿cómo hacemos para que al mismo tiempo que pone el fondo del pasto [señalar viñeta], escriba la palabra “HOLA”?

¿Qué bloque tengo que poner para que escriba “HOLA”? (¿qué bloque habíamos usado antes para que escriba “HOLA”?). Éste bloque [señalar en pantalla], muy bien.

Para que aparezca al mismo tiempo que el fondo del pasto, lo acomodo justo al lado.

Pero, si lo dejo así, ¿qué va a escribir? [señalar texto “abc” por defecto]. Entonces, ¿qué pongo acá? [señalar bloque de argumento]. Voy con la flechita y escribo “HOLA”.

¿Listo? ¿Probamos? [pulsar bloque inicio]

¿Qué pasó? ¿Se fue el fondo del pasto [señalar viñeta]? Sí, ¿pero se fue la palabra “HOLA”?

¿Qué bloque usamos para decirle a la compu que escriba en blanco? Usamos el mismo bloque [colocar al lado] y borramos lo que dice abajo [borrar “abc” por defecto].

A ver…

[13]

¿Probamos de nuevo? [pulsar bloque de inicio]

¿Funcionó? Primero, ¿puso el fondo del pasto y al mismo tiempo la palabra “HOLA” [señalar viñeta]? Después, ¿puso el fondo vacío y escribió en blanco [señalar viñeta]? ¡Muy bien!

Para seguir, borramos lo que hicimos [señalar en cuaderno]: agarro el programa y lo llevo acá abajo.

[14]

Además, la compu puede hacer todos estos sonidos [señalar en cuaderno]. Y cada sonido tiene su número, ¿ven? [señalar dos primeras filas de números] Este bloque [señalar en cuaderno] le dice a la compu qué sonido tiene que hacer.

[15]

Entonces, ¿cómo hacemos para que la compu haga el sonido de las burbujas [señalar viñeta]?

¿Qué bloque ponemos primero siempre? El bloque de inicio, muy bien. [colocar]

Ahora, ¿dónde está el bloque que le dice a la compu que haga un sonido? Lo agarro y lo engancho al lado.

Pero, ¿este sonido [señalar en cuaderno] es el sonido número 1 [señalar bloque de argumento]? ¿Cuál es el número del sonido de las burbujas? ¿Nos fijamos? [volver al catálogo de sonidos] El sonido de las burbujas es el número 6 [señalar en cuaderno]

Entonces, ¿qué número tenemos que poner acá [señalar bloque de argumento]? Voy con la flechita y escribo “6”.

A ver…

[16]

¿Probamos? [pulsar bloque inicio]

¿Funcionó? ¿Hizo el sonido de las burbujas [señalar viñeta]? ¡Muy bien!

[17]

Ahora, ¿cómo hacemos para que al mismo tiempo que hace el sonido de las burbujas [señalar en viñeta], ponga el fondo de abajo del mar [señalar en viñeta]?

¿Qué bloque pongo para que muestre este fondo? [colocar al final] ¿Y qué número es el fondo de abajo del mar? Nos fijamos [volver al catálogo de imágenes]. Es el número 8 [señalar en cuaderno]. Voy con la flechita acá, y escribo “8”.

¿Probamos? [pulsar inicio]

¿El fondo apareció al mismo tiempo [señalar en viñeta] que el sonido de las burbujas [señalar en viñeta]? ¿O apareció después [señalar viñeta siguiente]? Apareció después.

Cuando hace un sonido [señalar en pantalla], la compu siempre espera a que el sonido termine antes de seguir [señalar en pantalla bloque siguiente]. Entonces, para que el fondo aparezca al mismo tiempo que el sonido, no podemos ponerlo después [señalar en pantalla].

[18]

Para que el fondo aparezca [señalar en viñeta] al mismo tiempo que el sonido [señalar en viñeta], tenemos que ponerlo antes [señalar en cuaderno].

Entonces, agarramos este bloque [bloque colocado al final] y lo acomodamos acá [entre bloques].

Antes de probar de nuevo, aprieto el botón rojo [señalar en pantalla] para que la compu arranque otra vez de cero [pulsar rojo].

¿Probamos? [pulsar inicio] ¿Funcionó? ¿Apareció el fondo al mismo tiempo [señalar en viñeta] que el sonido [señalar en viñeta]? ¡Muy bien!

Para seguir, borramos lo que hicimos [señalar en cuaderno, y borrar], y aprieto el botón rojo [señalar en cuaderno] para que la compu arranque otra vez de cero [pulsar rojo].

[19]

Además, podemos decirle a la compu que haga girar [revelar robot] este muñequito.

El muñequito siempre empieza acá arriba [señalar en cuaderno]. El muñequito siempre empieza acá arriba [señalar en robot].

Entonces, cómo hago para pedirle a la compu que haga girar el muñequito hacia allá [seguir flecha en cuaderno], y que frene acá abajo [señalar segunda viñeta]. Que haga girar el muñequito hacia allá [señalar en robot], y que frene acá abajo [señalar en robot].

Este bloque le dice a la compu que gire el muñequito hacia allá [seguir flecha del bloque en cuaderno].

Primero siempre ponemos el bloque de inicio.

¿Qué bloque ponemos ahora [señalar en cuaderno]? ¿Dónde está? Lo agarro, y lo engancho al lado [colocar].

¿Probamos? ¡Muy bien! Pero no frenó. Para seguir, aprieto el botón rojo [señalar en cuaderno] para arrancar otra vez de cero.

¿Cómo hago para que frene solo?

[20]

Este bloque le dice a la compu que frene al muñequito.

¿Dónde está este bloque en la pantalla? Acá, muy bien [señalar]. Lo agarro y lo engancho acá, justo al lado.

¿Probamos? [pulsar bloque inicio]

¿Qué pasó? ¿Funcionó? ¿Giró el muñequito? A ver, probemos de nuevo. Prestemos atención acá [señalar el robot; pulsar bloque de inicio nuevamente]

¿Giró el muñequito? Pero, ¡giró muy muy poquito! ¡frenó muy pronto! Entonces, ¿qué tenemos que hacer antes de frenar? Tenemos que esperar un poquito.

Agarro el bloque para esperar y lo acomodo acá, antes de decirle al muñequito que frene.

¿Probamos de nuevo? [pulsar inicio]

¿Funcionó? ¿Giró el muñequito hacia allá [señalar viñeta], hacia allá [señalar robot]? Sí, ¿pero frenó acá abajo [señalar viñeta], acá abajo [señalar robot]? ¡Frenó muy tarde! Esperamos mucho. Tenemos que esperar un poquito menos.

Probemos con un número más chico. Por ejemplo… cuatro. Voy con la flechita acá y escribo “4”.

A ver…

[21]

¿Probamos?

Antes de probar de nuevo, aprieto el botón rojo [señalar en pantalla] para arrancar otra vez de cero.

Probamos de nuevo. ¿Funcionó? ¿Giró el muñequito hacia allá [señalar viñeta], hacia allá [señalar robot]? ¿Frenó acá abajo [señalar viñeta], acá abajo [señalar robot]? ¡Muy bien!

Para seguir aprieto el botón rojo [señalar en cuaderno] para arrancar otra vez de cero [pulsar rojo].

[22]

¿Cómo hacemos ahora para que el muñequito gire hacia el otro lado, hacia allá [seguir flecha en cuaderno]? Hacia allá [señalar en robot].

Este otro bloque le dice a la compu que gire el muñequito hacia allá [seguir flecha del bloque en cuaderno]. ¿Dónde está este bloque en la pantalla? Acá [señalar en pantalla], muy bien.

Entonces, queremos cambiar este bloque [señalar giro horario en pantalla], por este otro [señalar en pantalla].

Para sacar este bloque [señalar giro horario en pantalla], primero corro estos dos [señalar en pantalla] un ratito [correr].

Ahora, agarramos éste [señalar en pantalla] y lo llevamos acá abajo [borrar].

Ahora, ¿qué bloque queríamos poner? Este bloque [señalar en pantalla], muy bien. Lo agarro y lo engancho acá [colocar].

Por último, vuelvo a enganchar estos dos [señalar en pantalla y colocar].

¿Probamos? Funcionó? ¿Giró el muñequito hacia allá [señalar viñeta], hacia allá [señalar robot]? ¿Frenó acá abajo [señalar viñeta], acá abajo [señalar robot]? ¡Muy bien!

Ahora, vamos a ver unos ejemplos.

[cerrar modo *instrucción* e iniciar modo *práctica*; sustituir libro de instrucciones por libro de ejemplos]

Parte 2

[1]

Acá [señalar pantalla del reproductor] van a aparecer unos videos. Tengo que usar estos bloques [señalar panel inferior] para copiar lo que aparece en el video.

Aprieto este botón [señalar] para ver el video [reproducción]. [mientras se reproduce] Aprieto el mismo botón [pausa] para pararlo donde queramos. Y aprieto otra vez para seguir [reproducción].

[esperar al final]

Con este botón [señalar] lo vemos otra vez. Podemos verlo todas las veces que queramos. A ver, prestemos atención [recarga].

[pausar después de la trompeta]

Lo paro acá. ¿Qué es lo primero que tiene que pasar? Tiene que hacer el sonido de la trompeta [señalar la trompeta].

¿Qué bloque ponemos primero siempre? El bloque de inicio. [colocar]

Ahora, ¿qué bloque ponemos para que haga un sonido? Lo engancho al lado [colocar].

[presentar catálogo]

Pero, ¿el sonido de la trompeta [señalar la trompeta] es el número 1 [señalar bloque de argumento]? ¿Qué número es el sonido de la trompeta? Es el número 16 [señalar en catálogo]. Voy con la flechita acá y escribo “16”.

¿Seguimos? Aprieto acá [señalar] para seguir. A ver, prestemos atención [reproducción].

[pausar en viñeta siguiente]

Lo paro acá. ¿Qué tiene que pasar después? Tiene que mostrar este fondo [señalar] y al mismo tiempo tiene que escribir “HOLA”.

¿Qué bloque ponemos para que muestre un fondo? Lo engancho al lado.

[voltear catálogo]

Pero, ¿este fondo [señalar] es el número 1 [señalar bloque de argumento]? ¿Qué número es este fondo [señalar]? Es el número 19 [señalar en catálogo]. Voy con la flechita acá y escribo “19”.

Ahora, ¿qué bloque ponemos para que escriba “HOLA”? Lo engancho al lado.

Pero, si lo dejo así, ¿qué va a escribir? [señalar texto “abc” por defecto]. Entonces, voy con la flechita acá y escribo… escribo “HOLA”.

¿Seguimos? A ver... [reproducción].

[pausar en viñeta siguiente]

Lo paro acá. ¿Qué tiene que pasar después? Tiene que mostrar el fondo vacío y escribir en blanco [señalar]

¿Qué bloque ponemos para que muestre el fondo vacío? Lo engancho al lado. Y para que muestre el fondo vacío, voy con la flechita acá ¿y qué hago? Borro lo que dice acá abajo

Ahora, ¿qué bloque ponemos para que escriba en blanco? Lo engancho al lado. Y para que escriba en blanco, voy con la flechita acá ¿y qué hago? Borro lo que dice acá abajo, muy bien.

Pero, ¿qué tenemos que hacer antes de poner el fondo vacío y escribir en blanco [señalar ambos bloques]? Tenemos que esperar. Agarro el bloque de esperar y lo acomodo acá, en el medio, antes de poner el fondo vacío y escribir en blanco [señalar ambos bloques].

¿Seguimos? A ver… [reproducción]

¡Terminó! ¿Lo vemos otra vez todo seguido? A ver, prestemos atención. [recarga]

[esperar a que el video finalice]

Ahora con el bloque de inicio probemos lo que hicimos a ver si pudimos copiar el video. A ver…

¿Funcionó?

Primero, ¿hizo el sonido de la trompeta? Después, ¿mostró el fondo y escribió “HOLA” al mismo tiempo? Al final, ¿mostró el fondo vacío y escribió en blanco?

Entonces, ¿pudimos copiar el video? Parece que sí, muy bien. Antes de seguir, probemos de nuevo para estar seguros. Podemos cambiar lo que hicimos y volver a probar todas las veces que queramos.

Antes de probar de nuevo, aprieto el botón rojo [señalar en pantalla] para arrancar otra vez de cero.

A ver, probemos de nuevo…

¿Funcionó? ¿Pudimos copiar el video? ¡Muy bien!

[2]

Esta es la respuesta al ejemplo 1. ¡Muy bien!

Aprieto este botón [señalar] para pasar al ejemplo siguiente [siguiente].

[3]

A ver, prestemos atención al video [reproducción].

A ver, ¿lo vemos otra vez? [recarga].

[viñeta 1]

Lo paro acá. ¿Qué es lo primero que tiene que pasar? Tiene que mostrar el fondo del mar y al mismo tiempo tiene que hacer el sonido del agua.

¿Qué bloque ponemos primero siempre? El bloque de inicio. [colocar]

Después, ¿qué bloque ponemos para que muestre un fondo? Lo engancho al lado.

¿Y qué número es el fondo del mar [señalar fondo en pantalla]? Es el número 3 [señalar en catálogo]. Voy con la flechita acá y escribo “3”.

Ahora, ¿qué bloque ponemos para que haga un sonido? Lo engancho al lado [colocar].

[voltear catálogo]

¿Y qué número es el sonido del agua [señalar ícono del sonido en pantalla]? Es el número 8 [señalar en catálogo]. Voy con la flechita acá y escribo “8”.

¿Seguimos? A ver… [reproducción]

[viñeta 2]

Lo paro acá. ¿Qué tiene que pasar después? Tiene que girar el muñequito hacia allá [señalar sentido en pantalla].

¿Qué bloque ponemos para que gire el muñequito hacia allá [señalar sentido en pantalla]? Lo engancho al lado.

¿Seguimos? A ver… [reproducción]

[viñeta 3]

Lo paro acá. ¿Qué tiene que pasar después? Tiene que frenar el muñequito acá abajo [señalar], y al mismo tiempo tiene que mostrar el fondo vacío [señalar]

¿Qué bloque ponemos para frene el muñequito? Lo engancho al lado.

Pero, ¿qué tenemos que hacer antes de frenar el muñequito [señalar espacio entre bloques]? Tenemos que esperar [colocar].

Pero para que frene acá abajo [señalar], ¿10 [señalar bloque de argumento] está bien o es mucho? Es mucho, tenemos que poner un número más chico o nos vamos a pasar. Voy con la flechita acá y pruebo con… pruebo con “4”.

Y después, ¿qué bloque ponemos para que muestre el fondo vacío [señalar]? Lo engancho al lado. Y para que muestre el fondo vacío, voy con la flechita acá ¿y qué hago? Borro lo que dice acá abajo, muy bien.

¿Seguimos? A ver… [reproducción]

¡Terminó! ¿Lo vemos otra vez todo seguido? A ver, prestemos atención.[recarga]

[esperar a que finalice el video]

Ahora probemos lo que hicimos a ver si pudimos copiar el video. A ver…

¿Funcionó?

Primero, ¿mostró el fondo del mar e hizo el sonido del agua al mismo tiempo? Después, ¿giró el muñequito hacia allá [señalar en robot]? Al final, ¿frenó el muñequito acá abajo [señalar en robot] y mostró el fondo vacío al mismo tiempo?

Entonces, ¿pudimos copiar el video? Parece que sí, muy bien. Antes de seguir, probemos de nuevo para estar seguros. Recuerden que podemos cambiar lo que hicimos y volver a probar todas las veces que queramos.

Antes de probar de nuevo, aprieto el botón rojo [señalar en pantalla] para arrancar otra vez de cero.

A ver, probemos de nuevo…

¿Funcionó? ¿Pudimos copiar el video? ¡Muy bien!

[4]

Esta es la respuesta al ejemplo 2. ¡Muy bien!

El botón para pasar al ejemplo siguiente [señalar] ya no funciona. Quiere decir que no hay más ejemplos.

Así que por hoy terminamos. Pero otro día van a volver a jugar ustedes, ¿sí? Volvemos al salón.

Anexo 3.6: Guión de explicación

Discurso empleado en la etapa de Evaluación (ver apartado Etapa de Evaluación inicial *de la sección "3.2.1.5 Procedimiento"). Usamos un ejercicio de ejemplo (video sin eventos) para la explicación. Indicaciones no verbales se indican entre corchetes.*

Acá [señalar pantalla del reproductor] van a aparecer unos videos.

Tenés que usar estos bloques [señalar panel inferior] para copiar lo que aparece en los videos.

Vas a apretar este botón [señalar] para ver el video [reproducción].

[mientras se reproduce] Con el mismo botón [pausa] podés pararlo donde quieras, todas las veces que quieras.

Y lo apretás otra vez para seguir [reproducción].

[esperar al final]

Con este botón [señalar] lo ves de nuevo. Podés ver los videos todas las veces que quieras.

Tenés que copiar los videos lo más parecido que puedas, usando la menor cantidad de bloques posible.

Cuando termines con un video, apretá este botón [señalar] para ir al video siguiente [pulsar].

Acá te pregunta si pudiste copiar el video. Podés responder "sí", "más o menos" o "no" [señalar cada uno].

Si respondes "más o menos" o "no" [pulsar "no"], te pregunta si querés volver a probar el mismo juego, o si querés seguir al juego siguiente [mantener pulsado "volver" un segundo para que se vea].

Igual, en cualquier momento podés volver al juego anterior con este botón [señalar].

Podés avanzar [señalar] y retroceder [señalar] todas las veces que quieras.

Todo lo que hagas queda grabado. Entonces, si por ejemplo te trabás y querés seguir [señalar], después podés volver a revisarlo cuando quieras.

Te dejo esto [mostrar catálogo]; fijate que está impreso de los dos lados [mostrar ambas caras].

Yo ahora me voy. Vengo en un rato y me mostrás todo lo que hiciste.

Si terminás, podés revisar los juegos anteriores. Y si no, me podés ir a buscar. ¿Dale? Buenísimo.

Si en algún momento necesitás ir al baño o algo, andá directamente; no hace falta que me preguntes.

Anexo 3.7: Distancias de Levenshtein

Descripción detallada de la estrategia empleada para calcular la precisión de cada ejercicio cuantificando la distancia de Levenshtein entre secuencias construidas por les participantes y la secuencia correcta para cada ejercicio.

Precisión por ejercicio

Para cada ejercicio se obtuvo y analizó una captura de pantalla por segundo (ver sección “3.2.1.8 Análisis de datos”). Luego, usando la distancia de Levenshtein como se describió en el texto principal, se calculó un puntaje para cada secuencia presente en la captura, y el más alto se usó como puntaje de la captura. Finalmente, el puntaje de la última captura en un ejercicio se usó como índice de precisión de dicho ejercicio. Sin embargo, se observó que algunes participantes limpiaban el lienzo del editor antes de continuar con el ejercicio siguiente. Por lo tanto, en los casos en que la última captura se encontraba vacía, el último máximo local de puntaje se usó como precisión del ejercicio (en vez del puntaje de la última captura). Estos casos corresponden a 31 ejercicios (7,28% del total de ejercicios “intentados”), pertenecientes a 15 participantes.

Traducción ASCII

Para calcular la distancia de Levenshtein entre secuencias, los bloques de comando y de argumento se tradujeron en caracteres del código ASCII, siguiendo una tabla de traducción (ver “Anexo 3.8: Tabla de traducción distancias de Levenshtein”), de forma que un bloque = un carácter. Los bloques de argumento se tradujeron según su contenido; en la tabla de traducción se incluyen: (i) números 1 a 20 (correspondientes a las imágenes y los sonidos posibles), (ii) palabras entregadas como argumento en bloques de texto en programas correctos (y sus posibles errores de escritura), (iii) argumentos vacíos, y (iv) casos especiales (ej., omisión del bloque de argumento). Finalmente, bloques de argumento con valores no hallados en la tabla de traducción fueron traducidos como “*”.

Costos de inserción/remoción

Tratar los bloques de comando y sus bloques de argumento de forma separada permite coincidencias parciales: es decir, bloque de comando correcto con bloque de

argumento incorrecto, y viceversa. Sin embargo, esto implica que las operaciones (remociones e inserciones) sobre bloques de comando transitivos (es decir, que admiten bloques de argumento; ver sección “3.2.1.2 Plataforma de programación”) involucren dos pasos: uno para el bloque de comando y otro para su bloque de argumento, y por lo tanto las hace más costosas que las operaciones sobre bloques de comando intransitivos. Para evitar esto, usamos una distancia de Levenshtein con costos de inserción/remoción personalizados: bloque de comando intransitivo: 1, bloque de comando transitivo: .66, y bloque de argumento: .33. En consecuencia, las longitudes de secuencia *long(sec)* utilizadas para calcular el cociente *r* (ver apartado *Precisión* de la sección “3.2.1.6 Desempeño”) es la suma de los costos de inserción/remoción de los bloques en la secuencia.

Secuencias correctas equivalentes

Los bloques denominados “de acción” provocan que las variables del entorno programable se actualicen sin retener al intérprete (ver apartado *Tipos de bloques* de la sección “3.2.1.2 Plataforma de programación”). Como resultado, el orden de los bloques en una serie de bloques de acción no cambia el resultado final. Por ejemplo, si el objetivo es cambiar simultáneamente la imagen y el texto en la pantalla virtual del entorno programable, es lo mismo usar un bloque de imagen primero y uno de texto después, que hacerlo al revés. Por lo tanto, las secuencias de los participantes no se compararon con una sino con una serie de secuencias correctas con orden de bloques equivalente, y se usó la distancia más corta.

Puntaje orden-independiente

La distancia de Levenshtein evalúa la composición y el orden de las secuencias simultáneamente. Una secuencia con todos los bloques de la secuencia correcta pero en el orden incorrecto requiere algunas inserciones y remociones para llegar al orden correcto. Por ejemplo, una secuencia con un bloque correcto en una posición equivocada estaría a una distancia d=2 de la respuesta correcta: una remoción de la posición incorrecta, y una inserción en la correcta. Pero esta secuencia tendría el mismo cociente *r* que otra secuencia con un bloque incorrecto. Es decir, cuesta lo mismo mover un bloque de lugar, que eliminarlo e insertar otro. Para evaluar la composición en bloques independientemente del orden de los bloques, se midió una segunda distancia entre las secuencias ordenadas alfabéticamente. Finalmente, se

calculó un puntaje combinado como el promedio entre el cociente *r* de ambas distancias.

Anexo 3.8: Tabla de traducción distancias de Levenshtein

Bloques de comando		
Bloque de comando	**Traduc.**	**Comentarios**
Inicio	S	
Rotar horario	R	
Rotar antihorario	L	
Detener	H	
Sonido	P	
Texto	T	
Espera	W	
Imagen	B	

Bloques de argumento			
Bloque de comando transit.	**Contenido del bloque de arg.**	**Traduc.**	**Comentarios**
Imagen	0	_	Argumento 0 o vacío producen mismo resultado con bloque de imagen
Espera	*n, con n > 0*	w	
cualquiera	0	0	
cualquiera	1	1	
cualquiera	2	2	
cualquiera	3	3	
cualquiera	4	4	
cualquiera	5	5	
cualquiera	6	6	
cualquiera	7	7	
cualquiera	8	8	
cualquiera	9	9	
cualquiera	10	a	
cualquiera	11	b	
cualquiera	12	c	
cualquiera	13	d	
cualquiera	14	e	
cualquiera	15	f	
cualquiera	16	g	
cualquiera	17	h	
cualquiera	18	I	
cualquiera	19	j	
cualquiera	20	k	
cualquiera		_	bloque de argumento presente, pero vacío
cualquiera	*ausente*	-	bloque de argumento ausente
cualquiera	HOLA	l	argumento correcto del bloque de texto en ejercicio 1, evaluación
cualquiera	OLA	l	error posible para “HOLA”
cualquiera	H0LA	l	error posible para “HOLA”
cualquiera	0LA	l	error posible para “HOLA”
cualquiera	HOIA	l	error posible para “HOLA”
cualquiera	OIA	l	error posible para “HOLA”
cualquiera	H0IA	l	error posible para “HOLA”
cualquiera	0IA	l	error posible para “HOLA”
cualquiera	PATO	m	argumento correcto del bloque de texto en ejercicio 3, evaluación
cualquiera	PAT0	m	error posible para “PATO”
cualquiera	RANA	n	argumento correcto del bloque de texto en ejercicio 4, evaluación
cualquiera	HEY	o	argumento correcto del bloque de texto en ejercicio 5, evaluación
cualquiera	HEI	o	error posible para “HEY”
cualquiera	EY	o	error posible para “HEY”
cualquiera	EI	o	error posible para “HEY”
cualquiera	LEÓN	p	argumento correcto del bloque de texto en ejercicio 6, evaluación
cualquiera	LEON	p	error posible para “LEÓN”
cualquiera	LE0N	p	error posible para “LEÓN”
cualquiera	LEO´N	p	error posible para “LEÓN”
cualquiera	IEÓN	p	error posible para “LEÓN”

Bloques de argumento			
Bloque de comando transit.	**Contenido del bloque de arg.**	**Traduc.**	**Comentarios**
cualquiera	IEON	p	error posible para “LEÓN”
cualquiera	IE0N	p	error posible para “LEÓN”
cualquiera	IEO´N	p	error posible para “LEÓN”
cualquiera	321	q	argumento correcto del bloque de texto en ejercicio 7, evaluación
cualquiera	YA	r	argumento correcto del bloque de texto en ejercicio 7, evaluación
cualquiera	EL	s	argumento correcto del bloque de texto en ejercicio 8, evaluación
cualquiera	EI	s	error posible para “EL”
cualquiera	LICEO	t	argumento correcto del bloque de texto en ejercicio 8, evaluación
cualquiera	LISEO	t	error posible para “LICEO”
cualquiera	LICE0	t	error posible para “LICEO”
cualquiera	LISE0	t	error posible para “LICEO”
cualquiera	LICE	t	error posible para “LICEO”
cualquiera	LISE	t	error posible para “LICEO”
cualquiera	LIC…	t	“LICEO” no encaja y fue identificado como “LIC...” por OpenCV
cualquiera	LIS…	t	“LISEO” no encaja y fue identificado como “LIS…” por OpenCV
cualquiera	ACHÚ	u	argumento correcto del bloque de texto en ejercicio 10, evaluación
cualquiera	ACHU	u	error posible para “ACHÚ”
cualquiera	SALUT	v	argumento correcto del bloque de texto en ejercicio 1, reevaluación
cualquiera	SALU	v	error posible para “SALUT”
cualquiera	SAL…	v	“SALUT” no encaja y fue identificado como “SAL…” por OpenCV
cualquiera	RING	x	argumento correcto del bloque de texto en ejercicio 3, reevaluación
cualquiera	RIN	x	error posible para “RING”
cualquiera	RIMG	x	error posible para “RING”
cualquiera	RIM	x	error posible para “RING”
cualquiera	AUTO	y	argumento correcto del bloque de texto en ejercicio 4, reevaluación
cualquiera	AUT0	y	error posible para “AUTO”
cualquiera	UY	z	argumento correcto del bloque de texto en ejercicio 5, reevaluación
cualquiera	UI	z	error posible para “UI”
cualquiera	HUY	z	error posible para “UI”
cualquiera	HUI	z	error posible para “UI”
cualquiera	AVES	A	argumento correcto del bloque de texto en ejercicio 6, reevaluación
cualquiera	ABES	A	error posible para “AVES”
cualquiera	LA	C	argumento correcto del bloque de texto en ejercicio 7, reevaluación
cualquiera	CLASE	D	argumento correcto del bloque de texto en ejercicio 7, reevaluación
cualquiera	CLAS	D	error posible para “CLASE”
cualquiera	CLA…	D	“CLASE” no encaja y fue identificado como “CLA…” por OpenCV
cualquiera	UN	E	argumento correcto del bloque de texto en ejercicio 8, reevaluación
cualquiera	OSO	F	argumento correcto del bloque de texto en ejercicio 8, reevaluación
cualquiera	OS0	F	error posible para “OSO”
cualquiera	0SO	F	error posible para “OSO”
cualquiera	0S0	F	error posible para “OSO”
cualquiera	BRUUM	G	argumento correcto del bloque de texto en ejercicio 10, reevaluación
cualquiera	BRUM	G	error posible para “BRUUM”
cualquiera	BRUUN	G	error posible para “BRUUM”
cualquiera	BRUN	G	error posible para “BRUUM”
cualquiera	BRUU	G	error posible para “BRUUM”
cualquiera	BR…	G	“BRUUM” no encaja y fue identificado como “BR…” por OpenCV
cualquiera	*sin listar*	*	

Anexo 4.1: Guión de instrucción

Guión empleado por le adulte experte en la etapa de Instrucción de les participantes en la condición de *adulto-instructor*, siguiendo las pantallas de la guía de instrucción (ver Figura 4.8) provista por la sección Instrucción del modo de funcionamiento “retest” de la plataforma experimental (ver sección “4.3.1.5 Plataforma experimental”).

Pantalla 1: Presentación (Figura 4.8A)

Les voy a enseñar un juego que se llama “set”.

En el juego hay cartas y estas cartas tienen dibujitos.

Los dibujitos pueden ser de distintas FORMAS (pueden ser diamantes, óvalos o viboritas), pueden ser de distinto COLOR (pueden ser azules, rojos o verdes), pueden tener distinto RELLENO (pueden estar así vacío -señalar carta y abstracción-, así rayado, o así todo pintado), y las cartas pueden tener distinta CANTIDAD de dibujitos (pueden tener uno, dos o tres dibujitos).

Entonces, tenemos que descubrir si las cartas forman un “set”, o no.

Para formar un “set”, cada característica (forma, color, relleno y cantidad) tiene que cumplir una regla. Esa regla dice que tienen que ser o todas iguales, o todas distintas.

Si todas las características cumplen la regla, las cartas forman un “set”.

Pero, si alguna característica no cumple la regla -o sea, si en alguna característica no son ni todas iguales ni todas distintas-, entonces las cartas ya no forman un “set”.

Pantalla 2: Ejemplo 1 (Figura 4.8B)

Tenemos que responder si estas tres cartas forman un “set” o no.

Empecemos con la forma. ¿Se cumple la regla “todas iguales o todas distintas? Sí, porque tienen todas la misma forma: son todas óvalos. Entonces, **para la forma** se cumple la regla.

Sigamos con el color. ¿Se cumple la regla? Sí, porque tienen todas el mismo color: son todas verdes. **Para el color** también se cumple la regla.

…Sí, porque tienen todas el mismo relleno: son todas así, todo pintado...

Entonces, como la regla se cumple para todas las características, estas tres cartas forman un “set”.

Pantalla 3: Ejemplo 2 (Figura 4.8C)

… *[Símil pantalla 2]*

Pantalla 4: Ejemplo 3 (Figura 4.8D)

Éste es un poco más difícil. Presten atención.

… *[Símil pantalla 2]*

Pantalla 5: Ejemplo 4 (Figura 4.8E)

Ahora veamos este último ejemplo.

¿Se cumple la regla “todas iguales o todas distintas” para la forma? Sí, porque son todas distintas.

Y para el color, ¿se cumple la regla? ¿Son todas iguales o todas distintas? No son ni todas iguales ni todas distintas: hay dos iguales y una distinta. Entonces, para el color no se cumple la regla.

Como una de las características no cumple la regla, entonces estas tres cartas ya no forman un “set”.